

Moderation of abortion debate

The chairman of the US Republican Party is not the most familiar source of wisdom on abortion, but everybody will benefit from what Mr Lee Attwater has learned in this month's elections in the United States.

THE months ahead may see a marked attenuation of the public rows about the practice of legitimate abortion. That, at least, seems to be one of the portents of this month's elections in the United States, when New York elected its first black mayor, when governorships were up for grabs in states such as New Jersey and Virginia and when several candidates for less exalted offices discovered, against expectation, that a strong anti-abortion platform could be an electoral disadvantage. The shift was marked enough for Mr Lee Attwater, the chairman of the Republican National Party proclaim that his party is a tolerant party that welcomes a variety of opinions, even on matters as contentious as abortion. Most of the anti-abortionists disappointed at the polls appear to be Republicans.

That abortion is a contentious issue, and will remain so, requires no explanation. Whatever the resolution of the tediously familiar argument about the point at which human life begins (conception, implantation, the emergence of the neural crest or the like), there is no doubt that the majority of fetuses now aborted are alive by almost any criterion. That is why, where abortion is permitted (in the United States, abortion is the prerogative of state governments, but the Supreme Court could change that), explicit conditions must be satisfied before an abortion can be carried out. The general principle derives from the notion that if, at childbirth, there is a choice between the survival of the mother and of the child, the mother's life takes precedence. To be able to demonstrate, say by amniocentesis, that a fetus is genetically handicapped is not by itself a sufficient cause for legalized abortion; in principle, at least, it is necessary for a woman seeking an abortion also to demonstrate that the birth of such a child will be damaging to her, if not physically then psychologically.

Opponents of abortion are of three kinds: those who call abortion murder whatever the circumstances, those who complain that the conditions that must be satisfied before an abortion may be legally carried out are insufficiently stringent and those who complain that the preconditions are so loosely applied that legal abortion is elective abortion. The arguments for legalized abortion, in similar order of increasing moderation, are that a woman's body (and whatever is inside it) is at her own disposal (the feminist view), that unwanted children are genuinely damaging and that legal but regulated abor-

tion is preferable to illicit practice on a similar scale. In an ideal world, governments would seek to blunt unresolved argument by diminishing the scale of the problem by post-abortion counselling and easier access to birth-control clinics, but these measures are not popular with extremists. Until this month's elections, this seemed to matter to politicians. It may be less important now.

This journal has no brief to grind an axe on the abortion issue proper and does not do so, but there are several related issues that affect what happens in laboratories. In the United States, the Secretary of Health and Human Services has just extended a moratorium on the use of fetal tissue in transplants out of — it now seems — misplaced respect for anti-abortion forces (see *Nature* 342, 105; 1989). In Britain, the government's bill to regulate embryo research due to be published this week will be a vehicle for a parliamentary battle on abortion. But everywhere even those content with the legislation on abortion that may be in force should be concerned that its extreme opponents ('right-to-life' or 'abortion rights', as the case may be) are implacably opposed. It cannot much benefit grown-up societies that there should be apparently irreconcilable conflicts over the propriety of the law.

Despite a long history of bitter dispute, there is a chance that the two sides will be able to reach an accommodation with each other. That does not require that conflict should vanish overnight, but merely that people who are implacably opposed should learn that the other side has a case, but perhaps not the whole of what the extremists claim. Attwater may not be the most likely source of good sense on abortion, but his plea for tolerance deserves a hearing. □

Brain drain in flood

Migrations of skilled Europeans in the past few weeks are only the beginning. Employers should respond.

BY the end of 1989, the numbers of people who will have moved permanently from one country to another will be a record for recent times. Even before the East German frontier with the West was opened last week, some 50,000 people had made their way around it in the previous two months. That exceeds, but not by much, the numbers who will have left the Republic of Ireland (mostly for the



United States) in 1989. Almost as a direct consequence, Turkish guest-workers will be displaced from West Germany in numbers comparable with those of the ethnic Turks being expelled from Bulgaria. In present circumstances, ethnic Hungarians in Romania would be heading for Hungary if their government would let them go, but the Soviet Union may end this year having lost 500,000 people, mostly to the United States and Israel. And then there are the boat people from Vietnam, not to mention those who escape from Cambodia while they can.

It is natural to regard these huge migrations as welcome signs of the liberalization of frontiers all over the world, but even voluntary migration is never trouble-free. People who complain of the trauma of moving house within a single country should multiply their impressions by some large number, say ten, to allow for the difficulties of settling into an entirely unfamiliar environment. That is why the populations picking up their own roots to go elsewhere are naturally self-selected to include the most employable, which means those whose skills are technical and thus most portable.

There is worse to come. If Eastern Europe and the Soviet Union now keeps to the spirit of the Helsinki agreements, *perestroika* will have to be made to work to keep the skill it needs where it is needed. But even within the West, the growing conviction that the European Economic Community will yet become a true common market will mean that there are quite unpredictable movements of skilled people between member countries, which will further complicate the general drift of people to North America. What the British call the brain drain will become a general phenomenon.

How can the losers in this new labour market defend themselves? The only simple answer is that they will have to pay market prices in some sense or another. That does not necessarily mean that money salaries must everywhere become identical; working and living conditions are also important. Researchers and academics generally, but also people with skills in electronics and computer usage, are especially vulnerable in present circumstances. Wise employers, governments included, should begin planning now for the structural changes required to succeed in the coming competition for competence. □

Conduct unbecoming

Runners-up to science prizes receive no recognition but should know better than to complain.

ONE of the stipulations in Alfred Nobel's will, finalized a year before his death in 1895, was that the prizes to be awarded annually in his name should be awarded for discoveries made in the preceding year. Almost without exception, this stipulation is disregarded. But another, which limits to a maximum of three the number of people who may share a prize, is still honoured. Dominique Stéhelin, a French virologist, continues to argue (see page 329) that he should have been the third person

named for this year's prize in physiology or medicine, awarded to Michael Bishop and Harold Varmus.

As all agree, Stéhelin was the main pair of hands behind the experiments that first showed that the oncogenes of tumour viruses are stolen and corrupted versions of genes from the cells that the viruses infect. So much is acknowledged by the fact that he is the first author on the 1976 *Nature* paper that reported this discovery and which the Nobel assembly cite as the seminal piece of work. Nevertheless, only the head of the laboratory in which Stéhelin was a visiting scientist and the head of an associated research team have been awarded the prize.

Is this rough justice? With the absolute rule that all Nobel decisions are final and without appeal, it will be 50 years before the reasoning of the Nobel committees can be examined. It seems likely, however, that two questions that will have been asked are these: would Stéhelin, had he been working in any other laboratory, have been an author on the seminal paper; and would the paper have emerged from the Bishop/Varmus milieu even had Stéhelin not been there? It is likely that the answers were no and yes, respectively, considerably diminishing Stéhelin's case.

Nonetheless, there is no rule to stop the Nobel authorities from including Stéhelin had they so wished. Precedent was clearly against it, with the exclusion of Jocelyn Bell from the 1974 physics prize although it was she who spotted the pulsar for which, in particular, Antony Hewish was awarded his half of the prize, as a notable example. On the other hand, Georges Köhler, the visiting scientist who worked on the discovery of monoclonal antibody technology in César Milstein's laboratory, shared the 1984 prize for physiology or medicine. It is, of course, up to those who award prizes to set the rules but in the case of the relative merits of dirty hands to fertile minds there seems not to be one.

Stéhelin's immediate reaction was ungentlemanly. His claims can surely not have been overlooked. There is no doubt that it is harder to achieve international visibility in Lille than in San Francisco, but Stéhelin, who shared with Sydney Brenner and Walter Gehring the valuable and esteemed 1987 Jeantet prize for European biomedical scientists, was certainly in view from Stockholm.

But how wise is his new outburst, which seems largely to have been occasioned by some remarks by a member of the committee that recommended the winners? If the press has accurately reported these remarks, Stéhelin has some cause for complaint, but his response will do him no good. Even so, his contributions should be generously acknowledged at the prize-giving ceremonies in Stockholm next month. That will also be a time for the Nobel authorities to reconsider the terms of the prizes. With teamwork now the rule rather than the exception in research, for how long can they continue to award their prizes on the basis that the winners take all, and the runners-up receive no recognition? □

COBE starts its search for galactic fingerprints

- Satellite launched after four-year delay
- Looking for cosmic microwave irregularities

Washington

THE launch on 18 November, of the COBE (Cosmic Background Explorer) satellite from Vandenberg Air Force Base in California marks the start of a long-awaited astrophysical mission that had been delayed almost four years after the explosion of the space shuttle Challenger.

COBE's principal job is to look for irregularities in the cosmic microwave and far infrared backgrounds, an urgent task for cosmologists but one which has mostly passed beyond the capabilities of ground-based or balloon-flight experiments.

COBE was due to go on a shuttle flight soon after the Challenger disaster of January 1986, but NASA (National Aeronautics and Space Administration) decided shortly thereafter to switch it to a Delta rocket. This necessitated dismantling the finished satellite and putting the same instrumentation into a much smaller package. The reduction was achieved by omitting on-board thrusters, which would have been needed to propel COBE into its 900-km high, approximately polar orbit from the low shuttle orbit. The Delta launcher can put COBE directly into place.

The \$160-million COBE satellite carries three instruments, and will scan the entire sky twice during its one-year mission. By virtue of its polar orbit, COBE looks always at right angles to the direction of the Sun, whose light it must avoid to prevent damage to the detectors and interference with the measurements.

The two conical antennae of the Differential Microwave Radiometer will measure the difference in intensity of the cosmic microwave background, at three different frequencies, between two slightly displaced spots in the sky. This will yield an all-sky map of the variation of the microwave background from its average temperature of 2.7 K.

Temperature variations at some level are predicted by all theories of galaxy formation, the primordial variations in the cosmological density which acted as the 'seeds' of modern galaxies must also have left their mark as small irregularities in the microwave background.

Over the past two decades, experiments done from the ground, balloons and short-flight rockets have produced no conclusive detection of temperature variations. When the earliest experiments were done, simple galaxy formation theories predicted temperature variations of as much as one part in a thousand, but as increasingly

sophisticated instruments have brought the upper limit down to as little as a few parts in one hundred thousand, galaxy formation theories have been of necessity refined and adjusted to creep under the lowering wire of the measurements. There is a general feeling now that theory has few recourses left, and that if COBE does not find any temperature variations at a level perhaps an order magnitude less than the present best observations, then revision of standard cosmological ideas may be unavoidable.

The job of another of COBE's instruments, the Far Infrared Absolute Spectrophotometer, has become more crucial during the enforced delay of the satellite's launch. Last year, researchers from Nagoya University and the University of California at Berkeley published measurements obtained by a rocket-borne detector which indicated a substantial excess in the background intensity at wavelengths between 1.0 and 0.1 mm, above the peak frequency of the microwave background, but below infrared emission from interstellar dust. The Nagoya-Berkeley team saw emission in this far infrared region an order of magnitude more intense than they expected; no convincing theoretical source has yet been identified, and some cosmologists believe that the detection is the result of some systematic instrumental error. COBE's far infrared detector should confirm or invalidate the Nagoya-Berkeley measurements, in either case telling cosmologists something about a region of the spectrum which, because of atmospheric opacity, has been little studied so far.

COBE's third detector, the Diffuse Infrared Background Experiment, goes to even longer wavelengths, and will look for emission which could be the accumulated radiation from distant, early galaxies.

By scanning the sky over a frequency range from microwave to infrared, COBE is in effect scanning the process of galaxy formation, from the imprint of density fluctuations retained by the microwave background to the faint, red-shifted emission reaching us now from the first generation of protogalaxies. Cosmologists are guaranteed to get something to think about, especially from the microwave experiment: if COBE finds the long-sought irregularities, they will at last have some numbers to put into their theories; if it finds nothing, they will need some new theories.

David Lindley

PRIZES

Nobel dispute continues

London

DOMINIQUE Stéhelin, the French scientist who claims he should have shared this year's Nobel prize in physiology or medicine, has expounded his case in an open letter to the committee that recommended the winners, Michael Bishop and Harold Varmus. But the committee has no intention of responding.

In his letter, Stéhelin amplifies his view that it is unjust to have excluded from the prize the person who carried out the crucial experiments for the key discovery cited in the announcements of the awards. He argues that his exclusion is strongly detrimental because the "prize carries an unparalleled aura and prestige, lending it absolute and unquestioned authority in scientific quarters and the eyes of the public", and requests that the committee finds a way to repair the damage that has been done.

Apart from the fact that his name does not even appear in the documents that accompanied the announcement, most of Stéhelin's complaints are directed at comments made by one member of the committee in response to Stéhelin's initial outburst. Errling Norrby was quoted in newspapers as saying that the French researcher had not published anything of interest since 1976, when the key paper by Stéhelin, Varmus, Bishop and Peter Vogt was published in *Nature*, and that whereas Bishop and Varmus's contributions had already been recognized by several prize awards, Stéhelin had none.

Stéhelin, who directs a unit within the Institut Pasteur in Lille, responds that among the more than 100 papers he has published since 1976 are reports of the discovery of six new oncogenes, and that he has been awarded four prizes, among which is the 1987 Jeantet prize for European biomedical research, which he shared with Sydney Brenner and Walter Gehring.

Jan Lindsten, secretary of the Nobel committee that recommended Varmus and Bishop, says that Norrby was speaking in a personal capacity, and that the committee will not be responding to Stéhelin. It is, he said, increasingly difficult to select the appropriate winners as research is increasingly a collaborative affair. While pointing out that the committee is empowered to consider only those who have been nominated, he refused to say whether there were any nominations for Stéhelin. Lindsten emphasized that the prize was given for more than the 1976 paper alone and that the parts played by Stéhelin and many others had been considered by the committee.

Peter Newmark

Johns Hopkins as international host

Washington

WITH funds from the Howard Hughes Institute (HHI), researchers at Johns Hopkins University (JHU) in Baltimore, Maryland, are planning an international network of databases to support the mapping and sequencing of the human genome. The selection of JHU as host to the first database in the network puts a question mark over the future of the human genome mapping library in New Haven, Connecticut, which is directed by Frank Ruddle of Yale University, and has been supported by the HHI since 1985. But Purnell Choppin, president of HHI, said that researchers at JHU are also seeking federal funds and if they are successful the institute will not fund the database next year.

The new database at JHU, formally named the Genome Data Base (GDB), will provide access to primary data and literature citations, and daily updates of information generated by scientific workshops, the database's scientific editors and individual laboratories.

Users will also have access to OMIM, the on-line version of Mendelian Inheritance in Man, Victor McKusick's classic encyclopaedia of human genetic diseases. Peter Pearson, who worked at the University of Leiden in the Netherlands before he was hired by the HHI to be scientific director of the database, says that he was given a choice between Yale and JHU as the home of a new, more modern database, and that access to OMIM was one of the reasons he chose JHU. With access to both GDB and OMIM, users "will certainly get more for their money", he said. Other databases with which Pearson also hopes to forge links include the mouse genome database at Jackson Laboratory in Bar Harbor, Maine, and that at the Centre d'Etude du Polymorphisme Humain in Paris.

The second centre in the network will be created in London, with the collaboration of the Imperial Cancer Research Fund. Data from the GDB will be transferred at the international human gene mapping workshop in the United Kingdom, and at every subsequent workshop the data will be updated and transferred to a new centre. Eventually there will be an international network of human genome databases.

Kenneth Kidd of the school of medicine at Yale University, a consultant to the Yale database, says that it is trying to obtain funds from other sources, including the federal government, to replace the support from the HHI which this year was about \$1.2 million. The university was rejected as a candidate to host the new database amid complaints that the database there is in an outdated format, difficult to access and non-portable.

Christine McGourty

Japan presses on regardless

Tokyo

IN spite of the growing anti-nuclear, movement and the changing political environments, Japan is maintaining its commitment to expand nuclear power as its primary energy source, according to this year's white paper (policy document) on atomic energy, released last week by the Atomic Energy Commission (AEC).

Japan now has 37 commercial nuclear power plants in operation, supplying nearly 30 per cent of its electricity, and in its annual report two years ago the AEC outlined ambitious plans to increase that figure to 40 per cent by the year 2000. But in the latest report, such optimistic projections are conspicuously absent. The AEC notes that nuclear power has suffered setbacks outside Japan: in the United States, no new plants have been brought into operation for over a decade, West Germany has cancelled plans to build a nuclear-fuel reprocessing plant, and Sweden, Austria and Italy have voted against nuclear power.

Nevertheless, the Japanese report concludes that nuclear power is still "steadily expanding" worldwide and will remain central to Japan's energy policy.

But within Japan, opposition to nuclear power is growing (see *Nature* 333, 199; 1988). In the past 18 months, the government and power industry has poured tens of millions of dollars into advertising campaigns, lectures and a telephone service to try to win back public support for nuclear power. But surveys earlier this year show that opponents of nuclear power outnumber supporters two to one (see *Nature* 337, 494; 1989).

The government faces two severe tests of its nuclear power policy in the near future. A huge commercial complex to enrich uranium, reprocess spent nuclear fuel, and store low-level radioactive waste is scheduled to begin operations in 1991 in Aomori Prefecture, on the northern tip of the main island of Japan, but local farmers vehemently oppose the plan.

And in 1992, Japan hopes to start taking shipments of plutonium from Britain and France. The plutonium is derived from spent nuclear fuel from Japanese reactors, and is intended for use in advanced thermal reactors; Japan's first prototype fast-breeder reactor should become operational in 1992. But plans to ship the plutonium by air have already been abandoned because of opposition in the US Congress.

The AEC and the Science and Technology Agency (STA), which oversees Japan's nuclear power research and development, want to construct a new ship, to be operated by the Maritime Safety Agency, for plutonium transportation. But some members of the ruling Liberal Democratic Party (LDP) insist that exist-

ing ships of the Maritime Self-Defense Force should be used. However, despatch of these ships overseas is likely to cause public controversy, as it may violate military restrictions in Japan's post-war constitution. In addition, Japan has to face up to the problem of decommissioning nuclear reactors. Throughout the AEC report, nuclear power is referred to as "inexpensive" and "economical".

But the UK Central Electricity Generating Board, which has in the past promoted the low costs of nuclear power, had to admit at a recent public inquiry that the costs of decommissioning and waste disposal make nuclear power more expensive than coal power. This admission weighed heavily in the British decision two weeks ago to abandon plans to privatize the nuclear sector of the power industry (see *Nature* 342, 213; 1989).

However, British reactors are bulkier, older and more inefficient than their Japanese equivalents. Yukiko Araki, deputy director of STA's Office of Atomic Energy Policy Research, said at a briefing on the AEC report that decommissioning and waste disposal are expected to increase the cost of electricity from nuclear power plants in Japan only to ¥10 (7 cents) per kilowatt hour from the present rate of ¥9. The power industry recently established a fund derived from utility charges to cover the extra costs.

But finding a site for disposal of Japan's high-level waste is another headache for the government. Early plans for deep-sea disposal were abandoned because of objections from neighbouring Pacific island nations. A borehole survey of a site on Hokkaido, Japan's northern island, has been carried out and, says Araki, the government plans to bury high-level waste there at a depth of hundreds of metres. But Hokkaido residents oppose the plan and finding alternative sites will not be easy, especially because of the prevalence of volcanoes and earthquakes throughout Japan.

Finally, the recent rise in strength of the Japan Socialist Party (JSP), following a series of money and sex scandals that have forced the resignation of two prime ministers this year, has added a new political force to the fight against nuclear power.

The JSP has traditionally advocated the abolition of nuclear power, but now takes the more pragmatic position that it will not expand the industry if it gains power in the imminent general election.

Araki says that people in the industry do not now expect Japan to reach the commission's target of 40 per cent nuclear-power-derived electricity by 2000, and that the Ministry of International Trade and Industry is working on new targets.

David Swinbanks

Not just in California . . .

Washington

THE biggest earthquake of historical record in the United States occurred not on the West Coast but near the town of New Madrid, Missouri, about 200 miles north of Memphis, Tennessee.

During a period of seven weeks in 1811-12, three earthquakes, each of estimated magnitude greater than 8 on the Richter scale, altered the course of the Mississippi River, and were felt in Chicago and Washington. Although no very large earthquakes have hit the American midwest since that time, one of the political aftershocks of the recent Loma Prieta earthquake in California (see *Nature* 341, 676; 1989) has been an effusion of public concern over the dangers of earthquakes throughout the country. In order not to be accused of geographical chauvinism, Congress held a hearing last week to discuss the activity of the New Madrid fault zone, and the readiness of local authorities to cope with a catastrophe.

Unlike the geological instability of the west coast, which occurs at the boundary of different tectonic plates, faulting in the New Madrid region is caused by a mid-plate rift. As in the Mid-Atlantic and Red Sea rifts, subterranean material wells up as lateral forces pull the plate apart, but the New Madrid fault, according to Dallas Peck, director of the US Geological Survey, is now virtually static. The faults

are buried beneath alluvial deposits carried down by the Mississippi, making the system complex and unpredictable.

Empirically, faults between magnitudes 6 and 7 are estimated to occur about every 90 years. As recent events in California and Armenia have demonstrated, the difference between disruption and disaster from earthquakes of this magnitude has much more to do with building codes and education than geology (see *Nature* 337, 107; 1989). Arch Johnston, director of the Center for Earthquake Research and Information at Memphis State University, said at the hearing that progress on a local statute enforcing tougher building standards had accelerated remarkably after the Loma Prieta earthquake, and that new laws could be in place by the end of this year. But those measures apply only to the city of Memphis and its local county, and contain no requirements for retrofitting of existing buildings.

Although these issues and others, such as the availability and cost of earthquake insurance, were seriously discussed, it is unlikely that any real initiatives will emerge from last week's hearing. As Don Ritter (Republican, Pennsylvania) put it, "the 1812 New Madrid earthquake is probably as far from the thoughts of the citizens of Memphis, Tennessee, as the 1812 overture is from the citizens of Nashville".

David Lindley

LOMA PRIETA EARTHQUAKE

Stanford in squabble over relief funds

San Francisco

STILL reeling from the physical damage, estimated at \$160 million, suffered in the Loma Prieta earthquake on 17 October, officials at Stanford University, 40 miles south of San Francisco, found themselves on the public relations defensive last week after a senator in the California legislature maligned them for a lobbying campaign to get state relief funds that, it was suggested, were better deserved by more needy groups.

The controversy started when Democratic state senator Bill Lockyer tried to limit to \$1 million the amount of state aid a private, non-profit institution such as Stanford could receive. But he quickly ran into opposition as the package wound its way through the layers of state government. First, the limit on funds was raised to \$5 million. Then, when the bill was finally signed into law earlier this month by Governor George Deukmejian (Republican), the limit was waived altogether unless relief funds ran out.

The final form of the law upset Lockyer, who directed his anger at Stanford University. "Stanford was the main lobbying force arguing against any cap at all", he claimed,

adding that representatives of the university attended all the discussions: they were "in the common hearings, at the lobbying gate, visiting people's offices and button-holing them for the whole week". The senator was incensed that the wealthy institution used its considerable legislative weight to try and push through its package, especially as it has a \$1,800-million endowment. "I'm not a Stanford basher", Lockyer said, "I think it's a wonderful institution, and their contributions to our culture are enormous." What irked him, he said, was that "a lot of other people or organizations that are less fortunate don't have similar capacities to take care of their own problems". For their part, Stanford officials were stunned by the accusations, and have written to the editors of both the campus paper and a local newspaper defending their actions. "I'm really surprised that this has somehow become newsworthy", said spokesman Larry Horton. He acknowledged that the cap was the subject of a major debate in the Senate, but said that the university was only supporting what it thought was the most reasonable position.

Robert Buder

Zacher is new president

Munich

THE senate of the Max-Planck Society (MPS) last week selected Hans Zacher, now director of the Max-Planck Institute for foreign and international social law, as the organization's new president. Zacher, who was the only nominee presented to the senate, will begin a six-year term in mid-1990. He is the first president to emerge

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Zacher will take office in mid-1990.

from the humanities section of the MPS in its 41 years, and succeeds Heinz Staab, who returns after his six-year term to the Max-Planck Institute for Medical Research in Heidelberg.

With 62 institutes and an annual budget of DM1,240 million (\$670 million), MPS is the largest research organization outside the universities in West Germany. Only ten of the institutes are in the humanities, the rest covering biology and medicine, chemistry, physics and technology. Zacher, who is 61 and a pioneer of social law research for whom the institute for foreign and international social law was founded in 1980, now faces problems familiar to the heads of all science organizations in West Germany: an expected wave of retirements among directors late in the next decade, accompanied by the need to find positions for a flood of young researchers. Some MPS members worry that the organization of the society is not flexible enough to handle all the changes (see *Nature* 340, 335; 1989).

Steven Dickman

Market economy needed

Moscow

A transition to a regulated market economy, along with a large-scale decentralization of management, is the prescription given by the Soviet Academy of Sciences for the ailing Soviet economy.

These changes, and a number of urgently needed steps to bring them about, were recommended after a roundtable discussion held at the academy on 9 November, and attended by economists, People's Deputies (members of parliament) and officials from federal ministries and departments.

Many of the recommendations are aimed at 'westernizing' the consumer market: to bring money that is at present held as individual savings back into circulation, housing and land should be put up for sale, and businesses should offer shares to the public; more industrial manpower and resources should be allocated to the production of consumer goods and services; more consumer goods should be imported directly, instead of being made from imported raw materials. At the same time, the academy recommends, measures should be taken to restrain wage growth and restrict credit. With these changes, experts say, inflation will fall from its present rate of 8–10 per cent per year to 5–6 per cent, and consumer shortages will largely disappear in three years.

To support this new strategy of gearing the Soviet economy to the needs of the population, there must be stable investment in heavy industry, a reduction in the output of mines and a slashing of military spending: this will permit the present investment in consumer goods production and the service sector to be doubled, and still allow a 50 per cent increase in social spending.

But the authors of the academy report, which is to be submitted to the Supreme Soviet and to the government, say that it will be impossible to maintain a Western-style consumer goods market without complete reform of the pricing system and the creation of free markets in investment and credit, in manpower and in convertible currency. The present system of price-setting, in which the prices of fuels, raw materials and chemicals, for example, can be greater or less than the going world rate by as much as a factor of two, is fundamentally flawed, the report declares, and should be gradually changed. By 1995, the ratio between the numbers of products with fixed and market prices could be 30 to 70.

With price reform and a massive changeover to wholesale trade, the academy believes that the range of products in short supply would begin to shrink and most goods would then begin to circulate freely on the consumer market. But

some economists are more cautious, believing that a radically revised system can succeed only if shortages, and the consequent social tensions, are alleviated first. That problem is inseparably linked with the question of state subsidies for foodstuffs, which now total over 100,000 million rubles a year. Retail price reform would allow this figure to be reduced by as much as 75 per cent, but the population would have to be protected from the inevitable increases in foodstuff prices by comparable increases in income.

Concomitant with all this, a commercial banking system would have to be installed. The report foresees that a free money and investment market, coupled to a realistic foreign exchange rate, could take shape between 1993 and 1995.

But a genuine market economy has its own set of ailments, notably inflation, unemployment and bankruptcy. To safe-

NUCLEAR PROLIFERATION

Nuclear technology for Brazil

Munich

DESPITE opposition charges that Brazil has used West German nuclear technology to develop bomb-making capabilities, West Germany last week extended, for five years, an agreement to provide Brazil with more technology. The government withstood a fiery attack in the Bundestag (parliament) from the opposition Social Democrats (SPD) and Green Party, who claimed that the government was pandering to industry. Brazil is a significant importer of West German nuclear technology, which has lost its domestic market.

The government says there is no proof that Brazil can make nuclear weapons, and points to a clause in the Brazilian constitution that forbids Brazil from developing nuclear weapons.

Ulrich Irmer (Free Democrat) cited in an October debate in the Bundestag a report prepared by the West German intelligence service *Bundesnachrichtendienst*, according to which Brazil is not capable of producing weapons-grade uranium.

Nuclear reactors in Brazil are subject to oversight by the International Atomic Energy Agency in Vienna, and the West German government says it is satisfied with these various safeguards. But the Social Democrats (SPD), who have been pushing for tighter export restrictions in sensitive areas such as nuclear power and chemical weapons (see *Nature* 337, 678; 1989), say that government is not telling the whole story. Because Brazil has not signed the Nuclear Non-Proliferation Treaty (NPT), SPD Bundestag member

guard the population against these negative effects of reform, the academy proposes, for example, linking wages with some kind of regularly revised and published consumer price index. The future of *perestroika* depends, the economists assert, on the ability of the government to overcome the public's distrust of the state, which means that the state must give up its universal proprietary rights and concede the idea of diversified ownership.

The report criticizes draft Soviet legislation on leases, arguing that it constitutes an insufficiently radical reform of state property, and that it should be modified to allow unreservedly the rights of leaseholders to control income from a lease, acquire productive assets and hire labour, provided only that terms are agreed in advance. The academy also comments on a proposed new tax system, and says that the minimum taxable income should be higher and the maximum tax rate, now 80 per cent, lower. **Oleg Borisov**, Novosti

Edelgard Bulmahn argues that it may have a military programme intent on developing weapons. Only by signing the NPT, which would allow international monitoring of all its nuclear facilities, could Brazil give a clear signal that it does not intend to develop weapons.

The SPD cites the Brazilian military's so-called 'parallel programme' to develop nuclear technology as a threat, and says that a reorganization of Brazil's nuclear programme in 1988 legitimized the drive toward nuclear weapons. But Claus Jäger (Christian Democrat), who accused the opposition of trying to "stir up a scandal", says that government is "convinced" that nothing significant had changed. The original West German agreement, signed in 1975, was to provide up to eight nuclear reactors as well as enrichment and reprocessing technology to Brazil over a 15-year period.

Financial problems and technical hitches have led to a scaling-down, and Brazil now plans to build just three reactors.

Nevertheless, Brazil imports a considerable amount of nuclear technology from West German companies. Sources in the Bundestag estimated the value at DM 600 million (\$330 million) in 1989 alone, and the collapse of the market for nuclear power plants in West Germany, along with the decision not to reprocess spent nuclear fuel at home, makes the nuclear industry more dependent than ever on foreign markets.

Steven Dickman

■ For more on Brazil's nuclear capabilities, see page 374 of the supplement "Science in Brazil".

Indian zoologist suspected

New Delhi

A researcher at Bharatiyar University in the south of India is being investigated by the university for a deception allegedly practised on his British and US colleagues during scientific visits last summer. Professor G. Sundara Rajulu, head of the zoology department at Bharatiyar, has been accused by two US scientists of having 'perpetrated a hoax' by pretending that blood samples he collected in Britain came from a rare worm he claimed to have found in India.

The allegations were made in a letter written on 10 October to the chancellor of Bharatiyar University by Joseph Bonaventura, director of the marine biomedical centre at Duke University, North Carolina, and Charlotte Mangum, professor of biology at the College of William and Mary in Williamsburg, Virginia. Rajulu had spent two weeks as a guest in their laboratories during May and June.

The affair began when Rajulu contacted Bonaventura and Mangum, claiming that he had found haemocyanin, the copper-containing analogue of haemoglobin, in the rare Himalayan species *Eoperipatus Weldoni* a member of the phylum Onychophora. He wanted to bring some samples to the United States for further analysis. Onychophora are thought to be a link between annelid (legless) worms and arthropods, and the presence of haemocyanin could shed light on the link. Bonaventura and Mangum invited Rajulu to bring some live specimens to the United States for further analysis.

But when he reached their laboratories, Rajulu had no live specimens, and said they had been confiscated by US customs officials. Instead he produced a frozen blood sample and a preserved specimen of what he claimed was *E. Weldoni*.

The blood sample did indeed contain a haemocyanin typical of crustaceans, but on further examination turned out to have a general composition more characteristic of marine than land organisms. Mangum then discovered that Rajulu had come to the United States after a visit to Muriel Walker at the University of Leicester in England. An exchange of letters with Walker revealed not only that the blood sample Rajulu took to the United States was from a British shore crab, but also that he had told his UK hosts the same story about confiscated specimens that he had related to Bonaventura and Mangum.

Rajulu was supposed to have brought to Walker and her colleague Robert Harris live specimens of some other Himalayan creatures, but said that UK customs officials had confiscated all but one partial sample, which Harris later ascertained to be part of a common marine worm. The blood sample that Rajulu presented to

Bonaventura and Mangum was collected at his request by Harris.

Rajulu admits he has been served with an "explanation notice" by Bharatiyar University, but is surprised at the fraud allegation because, he says, Bonaventura and Mangum expressed no doubts about his discovery of haemocyanin during his visit to their laboratories. But on 26 July, after he had gone back to India, Rajulu sent Bonaventura a brief note saying that "in my anxiety to do things I committed a lot of mistakes" and apologizing for the

JAPAN

Huge profit from drugs

Tokyo

JAPAN'S hospitals and doctors are reaping enormous profits through sales of prescription drugs because pharmaceutical companies sell the drugs to medical professionals at a price far below the government's official reimbursement prices, an official of the Ministry of Health and Welfare admitted in the Diet earlier this month. The discounting practice has long been known to exist and is a major driving force behind excessive consumption of prescription drugs in Japan. But it is the first time a government official has put a figure on drug profiteering.

During a session of the auditing committee of the House of Representatives on 8 November, Tatsuhiko Sakamoto, director-general of the Health Insurance Bureau of the Ministry of Health and Welfare, under questioning by a member of the opposition Komeito party, revealed that the profits amount to about ¥1.3 million million (nearly \$10,000 million) per year. Medical institutions charged patients and the health insurance system about ¥5.15 million million yen for drugs in 1987 on basis of the ministry's official prices, but Sakamoto estimates the actual cost to the institutions to be only ¥3.83 million million.

Unlike medical practice in most developed nations where doctors are not allowed to sell drugs, nearly all hospitals and general practitioners in Japan have their own in-house pharmacies. To promote sales, pharmaceutical companies sell drugs to medical institutions at below the government's official price (the price that determines charges). There is thus a strong incentive for doctors to overprescribe and to prescribe drugs unnecessarily because they make a profit on every sale.

The profiteering and overprescription is a massive drain on the national health system and also has serious consequences for the health of the nation.

One example is antibiotics. In the West, doctors try to avoid prescription of anti-

bacterials because of the risk of development of antibiotic-resistant bacteria. But in Japan, antibiotics are liberally prescribed for all sorts of minor infections.

The Indian National Science Academy has also been sent a copy of the allegations but has yet to take any action. Bonaventura and Mangum agree that no great damage was done by the incident, apart from the waste of several researchers' time, but want to have the matter investigated because Rajulu is in the habit of arranging annual scientific visits outside India. Mangum says she knows of trips to Korea, Belgium and Vancouver, but has had no success in discovering if any similar deceptions occurred on those trips.

K.S. Jayaraman

otics because of the risk of development of antibiotic-resistant bacteria. But in Japan, antibiotics are liberally prescribed for all sorts of minor infections.

Another example is blood products. In the early 1980s, consumption in Japan surged with the import from the United States of cheap blood products which companies then sold to medical institutions at well below the government's official price. Among them were blood coagulants infected with AIDS, and most of Japan's 1,000 or so AIDS patients and carriers of the virus are haemophiliacs infected by the imported products.

For years, the ministry has been trying to separate the dispensing of drugs from doctors (a process called *bungyo* in Japanese). One approach has been to encourage the use of so-called 'legal prescriptions' which are handled by pharmacies outside the hospital. In response, large numbers of 'second pharmacies' have sprung up — although at first sight independent, they are in fact linked to medical practitioners. The doctor's wife, for example, may own the pharmacy.

Doctors and hospitals maintain that profits from drug sales are essential for management of their institutions. But slowly they are being brought 'reluctantly' towards full-scale *bungyo*, says Luke Gander of Pharma Japan, a newsletter for the pharmaceutical industry. But Gandor does not expect complete separation of doctors and pharmacies for 10 to 15 years.

There are too many vested interests to bring about rapid change. Pharmaceutical companies, both foreign and Japanese, are happy with their high sales in Japan. Doctors are happy with their profits. And the unquestioning Japanese public are happy to take home their bagfuls of unlabelled drugs after they visit the doctor. As the representative of one British pharmaceutical company observed, "It doesn't seem to do them any harm. The Japanese have the highest life expectancy in the world."

David Swinbanks

CF screening premature?

Washington

AT its annual meeting last week in Baltimore, Maryland, the American Society for Human Genetics (ASHG) issued a statement urging caution in the use of screening tests to alert couples to the presence of a genetic defect associated with cystic fibrosis (CF) and the consequent possibility that their children might be born with the disease. Geneticists agree that testing should be offered when either partner has a close relative affected with CF, but say they have "serious reservations" about widespread screening.

The race to develop a new test was prompted by the announcement in August of the cloning of the gene for CF by researchers at the University of Michigan, the University of Toronto and the Hospital for Sick Children, Toronto (see *Science* 245, 1059; 1989). Before then, CF could be detected only using conventional linkage analysis, which was available only to those with a family history of CF and with an affected relative still living. The discovery of the genetic mutation meant that a quick and cheap test could be developed based on a DNA probe specific for the mutation.

But the new test can still detect only about 70 per cent of CF carriers, because only that proportion of victims of CF have the identified mutation. Some researchers say that there should be no widespread screening for CF until the identification of additional mutations, which is expected within a year; others argue that the test is valuable despite its shortcomings and that widespread screening is worthwhile.

At its meeting last week, the ASHG HUMAN FRONTIER PROGRAM

Strasbourg office opens

Tokyo

THE long-awaited opening of the office for Japan's Human Frontier Science Program in Strasbourg, France was due to take place on Tuesday, 21 November (see *Nature* 340, 329; 1989). As of 17 November, more than 130 research groups led by scientists from the seven summit nations (Japan, United States, Canada, Britain, West Germany, France and Italy) and the European Communities had applied for the 20 or so \$500,000 grants that will be available, for research on the brain and molecular recognition, in the first year of full operation of the programme. Because of delays in opening the office in Strasbourg, the deadline for grant, fellowship and workshop applications will be extended to the end of this month, according to Toichi Sakata, director of the Frontier office at the Science and Technology Agency.

David Swinbanks

erred on the side of caution, saying general screening should wait until a larger proportion of CF carriers can be detected and also until there is a better idea of the education and counselling that could be given to CF carriers. To this end, says the ASHG, pilot screening programmes should be started as soon as possible. Peter Lanciano, the general manager of Integrated Genetics Inc., one of the companies providing a CF test, endorsed the ASHG statement, saying that it would do a disservice to the public to rush ahead with widespread screening "just because you can do part of the work".

He said that the company will refuse to provide the test to anyone who does not have a family history of CF. Frank Fujimura, scientific director of molecular biology at the Nichols Institute said that the institute would advise against taking the test but would not refuse it. But he admitted that resources are not available for providing counselling services on the scale required for routine widespread carrier screening for CF.

Fujimura said that although the laboratories involved are highly regarded, there is widespread concern at the lack of mechanisms to ensure that standards are maintained. In last week's statement, the ASHG also called for centralized quality control; both it and the College of American Pathologists are considering some kind of certification procedure for laboratories. The ASHG is now establishing a working group to study all the issues raised by the new test and produce a report next year.

Christine McGourty

BIOTECHNOLOGY

Celltech to change hands?

London

THE major shareholder in Celltech, the largest UK biotechnology company, is to sell its 36.4 per cent holding. It is considered likely that potential buyers will bid for all of Celltech's shares.

British and Commonwealth Holdings has been a shareholder from the start of Celltech in 1980. It has decided to sell its shares for "strategic reasons", and has asked the board of Celltech to identify possible buyers.

The news follows an announcement by chief executive Gerard Fairtlough, that he is to stand down from the company he has led from the start. Celltech, best known for its monoclonal antibody products, recorded its first profitable year in 1987. Last year was also marginally profitable.

Peter Newark

Limited use approved in India

Bangalore

INDIA'S National Monitoring Agency (NMA), set up by the government to study the problems associated with food irradiation, has approved limited irradiation tests with a view to general commercial use of the process, which kills bacteria that cause deterioration of food. Food irradiation became controversial in India following public protests over the import by India of Irish butter-oil that was alleged to have been contaminated by fallout from the Chernobyl nuclear explosion.

But now M.R. Srinivasan, chairman of India's Atomic Energy Commission, says that the NMA has approved the irradiation of spices and frozen sea food for export, as well as the irradiation of onions for the home market. To further the prospects of commercialization, the Department of Atomic Energy has joined with Gujarat Agro Industries Corporation and the Spices Board of Kerala, and intends soon to set up gamma irradiation facilities in a number of places in India.

Radhakrishna Rao

SOUTH AFRICA

New CSIR president

Cape Town

DR Brian Clark, 41, has been appointed president-elect of the South African Council for Scientific and Industrial Research (CSIR). The appointment was announced earlier this month by the Minister of Trade, Industry and Tourism, Mr Kent Durr. Clark will succeed Dr Chris Carberts, the current CSIR president, in September next year. He will be the sixth and youngest president of the organization, which is the largest of South Africa's three major research councils.

Clark joined the council's National Physical Research Laboratory as a technician after leaving school in 1965, and completed both his undergraduate and postgraduate training at the University of Pretoria as a part-time student.

Clark's present post in the CSIR is executive head of research, development and implementation (RDI). In this capacity he has been responsible for leading last year's restructuring of the organization into 12 technologically orientated divisions in the hope of attracting more research contracts from the private sector and government departments (See *Nature* 332, 197; 1988). In the current financial year, he expects outside contracts to amount to R174 million, accounting for 55 per cent of the RDI's budget.

Clark said in a statement that he hoped "to manage the CSIR in such a way as to make a major contribution to the ongoing development of South Africa".

Michael Cherry

WEST GERMAN UNIVERSITIES

Battle for youth plan

Munich

AT the same time as West German education minister Jürgen Möllemann has become the favourite of science organizations, he has become the bane of the education ministers of the *Länder* (states) in a simmering territorial dispute. Now he faces an all-out battle to gain support for a programme to support young researchers that could shape research and teaching at West German universities for the next 30 years.

Education had traditionally been a rump ministry when Chancellor Helmut Kohl chose Möllemann (Free Democrat or FDP) for the post in March 1987. When the ministry was created in 1972, the weekly newspaper *Die Zeit* wrote that if West Germany needed separate ministries for research and education, it should also set up separate ministries for letter post and parcel post. But circumstances besides his own ambitious nature helped Möllemann to overcome these hindrances: the FDP saw in the post a chance to put its liberal programmes across, and the federal budget was beginning to fatten after several lean years.

Möllemann soon realized that there is little for an ambitious education minister to do. His official brief covers just a few areas, including funding of the Deutsche Forschungsgemeinschaft (DFG), and construction at universities. So he began to try to steal areas of authority from other jurisdictions, mainly the *Länder* and, to a lesser extent, the Research Ministry.

Möllemann's style is to find a topical issue, overcrowding at universities for example, which brought students onto the

streets a year ago in larger numbers than had been seen in 20 years (see *Nature* 336, 704; 1988), then identify himself with the issue in the press — in this case by visiting the Munich university at the height of the unrest, where he was pelted with tomatoes. Finally, he makes a proposal to solve the problem: combined federal and *Länder* investment to provide more teachers and facilities in overcrowded fields such as business studies.

Möllemann's *Staatssekretär* Fritz Schaumann lists a number of achievements of which the minister is most proud: an increase in federal aid for students from middle-income families; the establishment of 40–60 graduate colleges to encourage more rapid progression to the PhD level; support for more research at technical colleges; and big increases in university building funds and for DFG.

For the long-term health of university research and teaching, the most important challenge is the plan to support young researchers for up to ten years until large numbers of professorships become available (see *Nature* 341, 94; 1989). This could help to persuade the best young people to remain at universities instead of going into industry or business. Despite criticism from finance ministers and even some education ministers in the *Länder*, the plan still has a fair chance of being passed at a 21 December 'education summit' with Kohl and *Länder* politicians in Bonn.

But the government will be hard pressed to find the DM6,000 million (\$3,200 million) that Möllemann wants for the plan. Priorities in Bonn have shifted since the huge influx of East German refugees began this summer.

Another important task, according to Schaumann, will be smoothing the way for the integration of West German courses of study into the European Communities, a tricky problem complicated by the differences in norms among the *Länder*, which fear losing autonomy to Brussels as well as to Bonn. Möllemann's activities have disturbed the *Länder*, to whom education policy is one of the last remaining bastions of power in an increasingly centralized country.

"Compared to what he has promised, the results are meagre", complains Thomas Goppel of the Christian Social Union

(*Staatssekretär*) in the Bavarian Ministry of Science. "If he wants his programmes to succeed, he should discuss them with us first," says Goppel, "he achieves ten per cent of what he has promised and then sells it as success." And "Möllemann is only interested in headlines", says Uwe Knüpfer of the Social-Democrat-run science ministry in North Rhine-Westphalia: "he ignores issues where behind-the-scenes negotiations would be more helpful than publicity campaigns." Goppel and Knüpfer say that Möllemann has done more for his own career than for students and professors.

But Möllemann has admirers in the science organizations. DFG president Hubert Markl says the minister "should be given credit" for this year's budget increase; Markl's pleading fell on deaf ears in the parliament and the government until Möllemann made DFG an issue. And Dieter Simon, chairman of the *Wissenschaftsrat*, says Möllemann's concern for science has done a lot of good: "even though his suggestions are not always the best, the fact that he keeps making them is a big help." Even critics such as Goppel agree that Möllemann has given a higher education a "new valuation" in the public mind. But both supporters and critics say that the biggest test for Möllemann — gaining federal and *Länder* support for his ideas — lies ahead. He cannot afford to walk away empty-handed from the 21 December education summit or the lasting effect of his contribution will, in the words of Knüpfer, "vanish in a puff of smoke".

Steven Dickman

RESEARCH FUNDS

More money for UK science

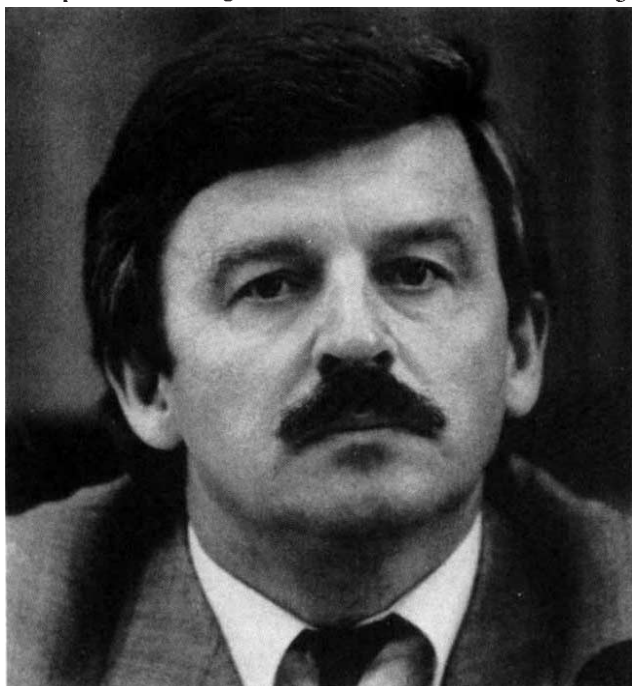
London

As part of UK Government expenditure plans announced last week, about £60 million will be added to the science budget for each of the next three years. Of the £61 million added to next years budget, £17 million is earmarked for the construction of a new research vessel for the British Antarctic Survey. Small sums are also allocated to the British Geological Survey, to enable it to set up a National Geological Information Service.

Next year's budget will now be £897 million, 27 per cent more than was spent in 1988–89. The Advisory Board for the Research Councils, which had been pushing for an even greater increase, will be deciding on how the extra money should be distributed among the research councils.

A simultaneous announcement of extra expenditure on higher education anticipates a predicted increase in the number of extra students. A system of student loans is to be introduced as part of the plan to boost numbers.

Peter Newark



West German education minister, Jürgen Möllemann.

Reasonable doubt?

SIR—George F. Wood's letter (*Nature* **341**, 100; 1989) suggests that some of your readers may not be aware of the legal profession's understanding of certain "issues involving probability". A substantial body of case law and commentary addresses the two issues he identifies. On the question of "what probability represents the 'reasonable doubt' beloved of the legal profession", see, for example, *Branion v. Gramly*, 855 F.2d 1257, 1263 n.5 (7th Cir. 1988) ("reasonable doubt means 0.9 or so, with adjustments depending on the gravity of the offense"); J. Kaplan *Stanford L. Rev.* **20**, 1065–1092 (1968). On the "purely statistical" proof of an elevation in "death rates due to . . . external circumstances", see, for example, *Sindell v. Abbott Laboratories*, 26 Cal. 3d 588, 163 Cal. Rptr 132, 607 P. 2d 924 (1980); D. Kaye *Am. Bar Fdn Res. J.* 487–516 (1982).

D. H. KAYE

Center for the Study of Law,
Science and Technology,
Arizona State University,
Tempe,
Arizona 85287-0604, USA

No paying twice

SIR—Henry Gee's News item (*Nature* **342**, 110; 1989) correctly explains the background and purpose of the programme of the European Science Foundation (ESF) on the molecular neurobiology of mental illness. ESF is asking for FF10 million (\$1.6 million) from its member organizations to run this programme for five years. Clearly this sum will not pay for all of the science involved; in each ESF country, the cooperating scientists will have to use normal national routes to raise the research funds required.

The ESF's FF10 million will be spent largely on coordination: interchange between scientists, collection of new pedigrees, computer services and the initiation of Southern blot analysis of chromosomal regions outside the normal interests of mental illness specialists, to ensure complete mapping of the genome.

There can of course be no question of anyone "paying twice" for the same work — no ESF member organization would do that.

Gee reports the concern of scientists about two points. First, I can reassure colleagues that the formal code of practice of the programme contains provisions for allotting particular regions of the genome to research groups that can justify a deep involvement with that region. Otherwise allotment will be

'random' to ensure both fairness and complete coverage of the genome. Second, the data facility of the programme will not perform any linkage analysis unless help is specifically requested by the members supplying the data.

We hope that this programme will help to identify innate risk factors in mental disease and provide a tool for the refinement of psychiatric diagnosis. These are high ambitions; to further them, we must of course satisfy both our professional colleagues and the proper concerns of funding organizations. We believe we are doing so.

JAQUES MALLET

CNRS Laboratoire de Neurobiologie
Cellulaire et Moléculaire,
Gif-sur-Yvette, France

Purge the bombers

SIR—The problem of terrorist bombings of airplanes in flight might be solved by a simpler and possibly more effective means than large, expensive thermal neutron analysers (TNAs). Specifically, the solution can make use of the elementary principle in chemistry that oxidation, including spontaneous combustion of plastic explosives, requires oxygen. Exclusion of oxygen, rather than or in addition to bombs, from airplane cargo holds could be accompanied by purging with nitrogen, an inert, completely non-toxic and non-polluting gas comprising nearly 80 per cent of the Earth's atmosphere. Perhaps more realistically, the large bins used to contain checked passenger luggage could be purged, and the purge maintained through slow release of nitrogen from a liquid or (small and safe) compressed source within each bin.

This proposal to exclude oxygen from luggage bins can be economically implemented in airports of even the most technologically non-advanced nations. The universal problem of flight delays would be turned into a benefit by providing time before airplane take-off for nitrogen substantially to replace oxygen within each piece of luggage. To function, a terrorist bomb would have to carry its own oxygen supply, or use air contained within hermetically sealed luggage to support an exclusion of sufficient intensity to breach the nitrogen-purged luggage bin, whereupon the explosion could accelerate by using air in the cargo hold, assuming it has not also been nitrogen-purged. It should be easy to scan luggage for hermetic sealing and/or the presence of compressed oxygen. Nonetheless, it would be desirable to identify any or all conditions under which single pieces of luggage could carry sufficient non-compressed oxygen to sustain explosions of the intensity required to breach nitrogen-purged luggage bins.

In airports equipped with TNAs, nitrogen-purging can address the problem

of unacceptably frequent false-positive TNA scans. Instead of manually searching numerous items of luggage yielding positive TNA scans, such luggage could be placed in nitrogen-purged bins, with some obvious trade-offs. As a final point, the feasibility of this low-technology solution can be tested in the short term, before major resources are committed to developing costlier high-technology solutions to be available only in the more distant future.

ROBERT A. MICHAELS

Ram Trac Corporation,
931 Northumberland Drive,
Schenectady,
New York 12309, USA

Nobel infallibility

SIR—The debate over the fairness or otherwise of the award of this year's Nobel Prize for Physiology or Medicine is reminiscent of the controversy over the 1952 prize for the discovery of streptomycin. This was awarded to Selman Waksman alone, specifically for his discovery of the antitubercular antibiotic streptomycin. Streptomycin was in fact discovered by one of Waksman's research students, Albert Schatz. Schatz had senior authorship on the main papers that announced the discovery and also appeared on the streptomycin patents. In addition, Waksman and others signed sworn affidavits to the effect that Schatz was co-discoverer. Finally, following a lawsuit, Schatz was legally defined as streptomycin's co-discoverer.

When Schatz's colleagues learned that Waksman alone was to be awarded the Nobel Prize for streptomycin, they began lobbying to correct the injustice. Some leading scientists, including William Feldman, sympathized with Schatz, but felt it politic not to intervene on his behalf. Amazingly, in reply to probing on Schatz's behalf, the committee that had awarded the prize admitted that it had never heard of Schatz. So much for the investigative abilities of the Nobel prize committees. It seems that nothing much has changed over the past 30 years or so. Indeed, the rules of the prize mean that no corrections can be made; essentially the Nobel committees regard themselves as infallible, and do not admit to mistakes. Stehelin should be warned that the scientific establishment does not take kindly to the Nobel prize boat being rocked. The career of Albert Schatz was blighted by his attempts to gain due recognition for his work. We can only hope that the same fate does not befall Stehelin.

M. WAINWRIGHT

Department of Molecular Biology
and Biotechnology,
University of Sheffield,
Sheffield, S10 2TN, UK

Popper on Schrödinger

SIR—I have read, somewhat unwillingly, P.W. Atkins' review of a biography of my late friend Erwin Schrödinger, which appeared under the title "Physicist and philanderer" (*Nature* 341, 193; 1989). I should be grateful if you would print this defence of a great man against the accusation of cowardice ("Schrödinger also appears to have been a coward . . .").

I do not know and I do not care who originated the accusation. But the accusation is absurd, as I will show. It illustrates how some people fail to understand the situation that prevailed in Germany and in Austria under Hitler.

Atkins writes of Schrödinger: "After fittingly succeeding Max Planck in Berlin in 1927, he (a Protestant) escaped to Oxford in 1933". This is a *suggestio falsi*: Schrödinger did not 'escape' from the Nazis, as did so many scientists (and others) who had been dismissed and who were threatened by a fate worse than dismissal. On the contrary, like Planck, von Laue and Heisenberg, Schrödinger was, in Hitlerian terminology, an 'Aryan', and like them he could have stayed in his position in Berlin and enjoyed wealth, power and influence. He did nothing of the sort. He gave up his position as the leading physicist in Berlin, and he chose, very soon after the dismissals had started in 1933, to join his 'non-Aryan' colleagues of his own free will. He did so with a public declaration expressing his disgust with Nazism.

Owing to the help of Professor Frederick Lindemann (later Lord Cherwell), he came to Oxford where I visited him in 1935 and 1936. I found him intensely unhappy. He was grateful to the powers that had made his reception at Oxford possible. But he was particularly unhappy about the scientific atmosphere there (or rather about the complete lack of it). He was not, of course, exclusively interested in physics: he was also very interested in classical scholarship. But he complained that, in Oxford, no serious discussion about scientific or scholarly problems was possible. It just was not 'done'. It was regarded as banalistic. Instead, one gossiped about people — usually running them down — and often rejected them in spite of their obvious competence and devotion to their subject. Schrödinger rejected this kind of snobbery (as did Churchill); and some silly people interpreted his open rejection (the openness of an apparent refugee) as ingratitude towards Oxford, and even towards England, which Schrödinger loved. That he ever "abused" Lindemann, or that he showed him "ingratitude", I regard as absolutely impossible. But Schrödinger was unused to dissembling, was highly temperamental and even a bit irascible.

And, as I know only too well from many discussions I had with him, he did not suffer fools gladly. But at the same time he was ready to make amends by showing special kindness. To say that he "considered himself to be God" is again absurd: he was very conscious of being far too frequently in the wrong, even morally.

During my visits to Oxford, when Schrödinger told me of his unhappiness, he also mentioned that he might leave Oxford and accept a chair in Austria which, in those days, was ruled by Schuschnigg who tried to keep the Nazis down. I implored Schrödinger to give up this mad idea. I told him that I was trying to leave Austria because Schuschnigg, who was after all a dictator like Mussolini and Hitler, was bound to succumb to Hitler. But Schrödinger thought that the allies would not let Hitler annex Austria — nor would Mussolini; and he said that, after all, if Hitler were to endanger Austria, he would emigrate again.

Not very long after this, I heard from Hans Thirring (my teacher in theoretical physics) that Schrödinger had declined Thirring's offer to vacate, in his favour, the chair in Vienna for a vacant chair in Graz, but that Schrödinger had himself accepted another chair in Graz.

In 1938, Hitler invaded Austria after a crisis of only a few days which left Schrödinger no time to escape (this time, the term 'escape' is appropriate). Being a known opponent of Nazism, he was at once in great danger. He was forced (like Galileo) to recant his anti-Nazi views; but being a little younger than Galileo was at the time, he managed an adventurous escape: he put a small suitcase with his most important documents into his old car and drove to the Italian frontier, where he left his car standing in a side road and crossed the frontier with his wife on foot. Once on the Italian side, he took a train to Rome. There he telephoned Dublin, trying to contact Eamonn De Valera, the Irish prime minister, who had invited him before. De Valera, he found out, was in Geneva. So he rang Geneva and, with some difficulty, reached him. He told him about his situation, and that his money was running out rapidly while he was telephoning. De Valera said he should at once take a train to Geneva. On the Italian-Swiss border, Schrödinger and his wife Annemarie were stopped by the *carabinieri*: they were forced to get out of the train, which left without them. They thought that the Nazis had alerted the

Italian fascists. They were both searched, but they were allowed to continue by a later train. (It turned out that because they had hardly any money on them, they had been suspected of trying to export money illegally.) Once across the border they were safe.

Some of my ('Aryan') friends in Austria managed to leave under tragic and dangerous circumstances. One of my closest friends, Otto Haas, a lieutenant in Hitler's army during the war, tried to organize a resistance movement and was hanged in 1944.

Some of my friends made their peace with Nazism. They were in danger for they had been *Judengenossen* (associates of Jews). Who is to blame them? Who is to call them cowards?

But Schrödinger was not one of them: he left Austria soon after its annexation under the most dangerous circumstances, even though he would have been welcomed by the Nazis had he recanted. Galileo did recant. Who dares to call him a coward?

A last remark on poetry, on *What is Life?*, and all that.

When in 1927 I first read Schrödinger's *Abhandlungen zur Wellenmechanik*, I was overwhelmed, not merely by the beauty of his theory, but by the sheer poetry of it: I knew then that he was a real poet. When, long after his death, I read some of his poems, I remembered, and felt confirmed in my opinion.

When I read, 40 years ago, *What is Life?*, I was greatly excited. But I soon realized that this was a kind of manifesto without much content; that Schrödinger the poet had run away with Schrödinger the scientist. After all, he was inspired by the views of Max Delbrück, and the book was meant as a popularization of Delbrück's ideas. (Schrödinger's own suggestion that what characterized a living organism was that it sucked up negentropy was not to be taken seriously, for every washing machine does that, without coming to life.) However, the book worked as a biologist's manifesto, and inspired many young people: no doubt, Schrödinger's poetic powers succeeded where, as I believe, his biology failed. Here, as everywhere else, I disagree with your reviewer. (But I largely agree with Linus Pauling, and more or less with Max Perutz, in their contributions to *Schrödinger*, edited by C. W. Kilmister, which was published by Cambridge University Press in 1987.)

A last word: gossip about Schrödinger's philandering is not my business; nor, I suggest, is it anybody else's business. And I find it hard to swallow that Annemarie Schrödinger is being made part of the gossip.

KARL POPPER

136 Welcomes Road,
Kenley, Surrey CR2 5HH, UK

Letters submitted for Correspondence should be typed, double-spaced, on one side of the paper only.

The nature of science

SIR—While A. Travis (*Nature* **341**, 10; 1989) may be bemused by the Universe, he is certainly befuddled about the nature of science. It would seem that, to Travis, only when a model can be “tested experimentally can it be held as truth”. I, and I am sure many other scientists, would take issue with this statement.

Science cannot and should not claim to be discovering truths. As fallible human beings and scientists, our best hope is to propose theories (or hypotheses) that are rigorous enough to withstand repeated testing and to be supported by the evidence that can be gathered. Many of the theories simply will not withstand the testing and will either be falsified outright or, more likely, be replaced by a new and better corroborated theory. This is the strength of the so-called ‘scientific method’ in attempting to understand the world around us. It cannot proclaim ultimate truths as Travis would suggest, for those who do so step out of the realm of science.

A more debatable issue is whether only experimental science is truly science. Both the aforementioned statement of Travis and a later comment that “untestable theories in science (such as the evolution of life through random diversification followed by natural selection) are permissible, but not if they involve forces from an unknowable source”, suggest that Travis would hold to the stricter ‘experimentalist’ perception of science. By most scientists’

perception of their profession, however, if a theory is untestable it is not science and is not permissible, at least within the realm of science. The statement only makes sense if one replaces “untestable” with “non-experimental”. If this is done, I have less quibble with the statement, because much of testable science is in fact non-experimental. Every science, but some more than others, has an historical component that represents a unique series of events that took place in the near or very distant past. The experiment has been run, but the evidence is there and can be corroborated over and over again just as an experiment can hopefully be repeated.

J. DAVID ARCHIBALD

*Biology Department,
San Diego State University,
San Diego, California 92182, USA*

SIR—A. Travis feels compelled to invoke a ‘creative force’ when considering the mysteries of the observable Universe, in particular those aspects that appear least amenable to scientific investigation. I doubt, however, whether it is appropriate to label such areas of mystery as being particularly ‘religious’. Such labelling is similar to that which became popular following the deistic movement of the early eighteenth century, in which ‘god’ or ‘religion’ were invoked to explain phenomena for which science had not yet found satisfactory theories. Not surprisingly, such a ‘god-of-the-gaps’ faded like the

smile on the Cheshire cat as science continued its advance. The invoking of a ‘creative force’ to fill the present gaps in our scientific knowledge does not seem any more satisfactory. The task of science is to construct better and better ‘maps’ that will incorporate an increasingly sophisticated understanding of the world around us, and there is no reason to suppose that topics such as ‘human intelligence’ and the ‘origin of life’ will not be much better defined in future maps than they are on the rather poor ones that we have at present.

Those who laid the foundations of modern science in the seventeenth century did not look for gaps in their scientific knowledge in which to make room for their religious beliefs. Quite the reverse: they saw behind the complete ‘map’ of scientific knowledge a God whose continuing creative activity made it a worthwhile enterprise to look for consistency and reproducibility in physical phenomena in the first place. Their sense of awe at the world around them came not so much from its unexplained mysteries as from the fact that it was explicable, and could be expressed particularly in the language of mathematics. For example, Galileo saw God as the “divine mathematician”, who had made a universe that was intelligible by the human mind, so that investigating “the true constitution of the universe” was “the most important and most admirable problem that there is”. Three hundred years later, our investigations of the ‘true constitution of the universe’ have a very long way to go, but it is good to emphasize that religion, at least the kind of religion that Galileo espoused, still has no hidden investments in scientific ignorance. The role of religion and philosophy is to address the much broader questions of whether the very existence of the ‘map’ itself, with ourselves as part of it, has any particular meaning in a more ultimate sense, and also to tackle the vital ethical questions as to how our scientific knowledge should be used. Peter Medawar has recently reminded us that the scientific method itself is not equipped to provide answers for such questions². But as we attempt to tackle them at other levels, the proposal of ‘unknown, creative forces’ as religious-sounding plugs for the gaps in our present scientific knowledge is as inappropriate for religion as it is for science.

DENIS R. ALEXANDER

*Cell Surface Biochemistry Laboratory,
Imperial Cancer Research Fund
Laboratories,
PO Box 123,
Lincoln's Inn Fields,
London WC2A 3PX, UK*

Sahni visit denied

SIR—In normal circumstances, I would have ignored V. J. Gupta's statement (*Nature* **341**, 11–12; 1989) that my father, the late M. P. Sahni, had visited his (Gupta's) localities in 1964, as not being central to the main issues at stake in the matter of the ‘peripatetic fossils’—see the article by Talent (*Nature* **338**, 613–615; 1989) and those by others (*Nature* **341**, 13–16; 1989). But the present situation is far from normal and warrants that the truth be made known regarding M. P. Sahni's involvement in the matter. Sahni had an unblemished and distinguished record of meticulously conducted research spanning nearly half a century. It is therefore all the more distressing that his name has been dragged into a controversy after his death, and this has provoked several of his friends, ex-colleagues and well-wishers to ask me to set the record straight.

The facts are that Sahni attended a seminar at Srinagar (Jammu and Kashmir) in October 1964. He left Srinagar on 16 October, seen off at the airport by Dr S.K. Shah, one of the organizers of that seminar and now professor of geology at the University of Jammu. During his brief

halt, Sahni neither visited the graptolite localities nor did he accompany the post-seminar field excursion and I have documentary evidence to substantiate this.

The controversy over the ‘peripatetic fossils’ has caused intense introspection and soul-searching among Indian earth scientists in general and palaeontologists in particular. Open systems based on trust are fragile institutions prone to attack, but the strength of such systems derives from the internal reaction and response in the hour of crisis. By this yardstick, the national response has been spontaneous and extensive: exposure in the news media and the setting up of two major commissions of inquiry and other bodies of investigation. Introspection has led to action.

ASHOKI SAHNI

*Centre of Advanced Studies in Geology,
Panjab University,
Chandigarh – 160014, India*

■ In the article by U. K. Bassi (*Nature* **341**, 15–16; 1989) the author's institutional address was used in error. Although U.K. Bassi can be contacted at the Geological Survey of India, his personal address should have appeared in the article. □

1. Drake, S (transl.) *Discoveries and Opinions of Galileo* 97, (Doubleday-Anchor, Garden City, New York, 1957).
2. Medawar, P. *The Limits of Science* (Oxford University Press, 1985).

Disentangling the greenhouse

Continuing uncertainties about the working of the greenhouse enhancement give the lie to those demanding, at a meeting last week in the Netherlands, that the details of an international convention should be settled on the spot.

Melbourne

EVERYBODY seems to know that this is the part of the Earth's surface at which the Antarctic ozone hole reaches furthest north. While the observations of anomalous springtime ozone depletion outside the Antarctic Circle covers only a single season (this time last year), people seem convinced that this will be the beach summer of total block, of warmth without tan (see *Nature* **340**, 290–294; 1989). Australians do not lightly forgo such pleasures, which is why some wonder how long-lasting will be the resolve to keep ultraviolet away from skin.

But even in literate Australia, the confusion between the ozone hole and what is strictly the supposed enhancement of the normal greenhouse effect by atmospheric constituents opaque to infrared is as profound as elsewhere. There may be rough justice in that. The springtime Antarctic ozone hole, the reality of which seems well-attested by this season's observations, is at least a sign that a measurable property of the atmosphere is affected by the rate at which unwanted refrigerants are discharged to the atmosphere. If chlorofluorocarbons (CFCs) can affect the ozone layer, why should not the much more copious releases of other greenhouse gases, such as carbon dioxide, affect the climate?

The confusion is nevertheless a threat to good causes, some of which are also green. CFCs are greenhouse gases in their own right (and, molecule for molecule, are more effective than CO₂). But to a first approximation, ozone depletion, whatever its cause, does not affect the climate or even the weather. To a second approximation, of course, ozone depletion (and so CFCs) affect both climate and weather through the changed variation of temperature with altitude — the tropopause is lowered — but the consequences cannot easily be predicted quantitatively let alone disentangled from the data.

How can the distinction between the ozone hole and the enhancement of the greenhouse effect best be established as a part of general knowledge? Much might be done by the direct measurement of the spectrum of solar ultraviolet radiation reaching the surface of the Earth at different places and seasons. The most obvious practical impediments are those of calibration, which, in the absence of data from the past, must necessarily be abso-

lute. Even so, it is remarkable that so little has so far been said about the flux of ultraviolet light reaching the surface of the Earth in even recent years. Paradoxically, if it were shown that ozone depletion means extra ultraviolet, it would be easier to separate that from the direct enhancement of the greenhouse.

While there should by now be no doubt that the next practical step should be the negotiation of an international convention to restrict the release of all greenhouse gases, the more obdurate impediments to understanding and prediction remain as obdurate as ever. It is maddening that the big uncertainties seem to change only slowly with the passage of time. Two stand out. The climate models can incorporate what is called average cloudiness, but the effect of real clouds, with real edges, on the heat balance in the troposphere could be qualitatively different. In principle if not in magnitude, the uncertainties are like those of estimating the role of clouds in nuclear winter (see *Nature* **318**; 99; 1985). That must remain a problem.

The other big uncertainty, at least so far as the greenhouse enhancement due to CO₂ is concerned, remains that of knowing the fate of whatever fraction of the atmospheric content is dissolved in the oceans every year. Increased conversion to inorganic carbon would be benign, solution as bicarbonate in a steadily deepening layer of warm water above the oceanic thermocline would be the opposite. That issue has been endlessly discussed, but there are few who at this stage would put their hands on their hearts and say they know the quantitative truth.

Even relatively operational questions, that of the movement of CO₂ within the oceans for example, are unresolved. That is nicely illustrated by an attempt last year, by Peter G. Brewer and Catherine Goyet of the Woods Hole Oceanographic Institution and David Dyrssen of the Chalmers Institute of Technology at Gothenburg, to measure the northwards flux of CO₂ in the North Atlantic (*Science* **246**, 477; 1989). The North Atlantic is usually supposed to be a substantial sink for the gas, which is supposed to be carried northwards by oceanic currents. But measurements across the Florida Gulf and at three stations along the 25th parallel in mid-Atlantic lead to the conclusion that while there is a substantial northerly transport, that is almost but not quite offset by trans-

port in the opposite direction. The net northwards flux, at 260 million tonnes a year, is just over one per cent of the flux in either direction, or just under 5 per cent of the annual production of this greenhouse gas. This is much smaller than expected.

What can be the explanation? One is that the data may be insufficient. Three sets of measurements of CO₂ concentration with depth are not many, although the authors of the study have used previous survey data to interpolate between their stations. Another is that, while CO₂ will be released as the lower waters of the Gulf Stream reach the surface, it will be reabsorbed from the atmosphere as the Gulf Stream cools, reducing the partial pressure of the gas. Still another is that expectations are not yet fulfilled because there has not been enough time, since the onset of rapid greenhouse gas production half a century ago, for the deep waters of the North Atlantic to have travelled as far north.

In the circumstances, it must plainly be foolhardy to pretend just what the future course of the enhancement of the greenhouse effect will amount to. Better and more measurements will help, although none will be decisive. But if more were known of the variation of CO₂ concentration with latitude in the deep waters of the Atlantic, it might at least be possible to guess when transport into the North Atlantic atmosphere would become substantial. That would provide a means of telling just how quickly the international convention must be negotiated and brought into effect.

Questions such as these were plainly in the minds of some of those attending last week's meeting in the Netherlands, one of several meetings planned as preparation for the more formal diplomatic conferences to be held next year. Governments such as the British were pilloried, but unfairly, for having asked that the details of a convention — the date at which the further accumulation of greenhouse gases should be halted, for example — should await further study. That, of course, is entirely sensible. It is in nobody's interests that the economic disruption an international convention will certainly bring should be accelerated. It is not that there is nothing else to talk about, the demand on behalf of developing countries for fair shares in any quota system there may be for one thing.

John Maddox

Tolerance: a second mechanism

Anthony L. DeFranco

A CENTRAL problem for the immune system is distinguishing foreign invaders from the body's self-components. Recognition of the former concentrates the destructive mechanisms of the immune system on the infectious organism; mistaken recognition of the latter results in autoimmune disease. On page 385 of this issue¹, Goodnow and co-workers report that immunological tolerance to self can be induced in mature B cells, thereby giving us good reason to think that tolerance involves at least two principal mechanisms.

Discrimination between self and non-self can arise from a developmental switch in the response of lymphocytes to antigen: contact of antigen with the antigen receptor on immature B or T lymphocytes induces either death of the cell ('clonal deletion') or long-lasting unresponsiveness ('clonal anergy'). One problem with understanding how this mechanism could be fully responsible for immunological tolerance to self is that lymphocytes develop in limited anatomical locations (bone marrow for B cells and thymus for T cells), but once mature they circulate throughout the body. Thus, mature lymphocytes may encounter self-components that were not present at the sites of maturation, and that therefore could not have induced clonal deletion or anergy. An attractive solution to this puzzle is that contact with antigen may, under some circumstances, also induce unresponsiveness in mature lymphocytes, and Goodnow *et al.* provide perhaps the clearest demonstration to date of this form of tolerance induction.

We are currently in the third phase of the investigation of immunological tolerance. The first phase, ushered in by the studies of Owen *et al.*² with dizygotic twins of cows, established that tolerance is an acquired characteristic. The second phase involved analysis of tolerance at the cellular level. These experiments established that unresponsiveness was a property of the antigen-specific lymphocytes, that direct *in vitro* antigen exposure of B cells could induce unresponsiveness, and that clonal deletion or anergy was most easily achieved with immature B lymphocytes³. Less was known about tolerance of T cells, because it was harder to follow their activity. A great problem of the second phase was that antigen-specific lymphocytes could not be followed directly. In some experiments, antigen surrogates (such as anti-IgM antibodies) were used to circumvent this problem, but such an approach had inherent limitations.

With the advent of mice expressing transgene-encoded antigen receptors⁴⁻⁷, and with the discovery that antibodies

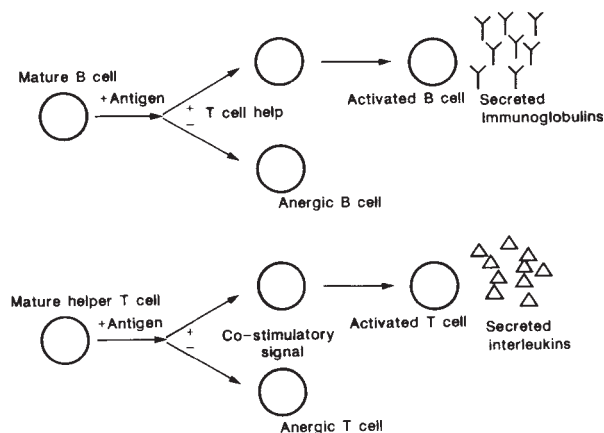
recognizing the products of certain gene segments (V β) can distinguish all T cells recognizing certain antigens⁸, it is now possible to look at lymphocytes of a given specificity and analyse the effects when they encounter antigen. A great deal has already been learned from these third-generation studies. When immature T cells developing in the thymus encounter antigen and appropriate proteins encoded by the major histocompatibility complex (MHC), the result is clonal deletion^{4,5,8}. Similarly, immature B cells expressing immunoglobulin transgenes are inactivated, either by clonal anergy⁶ or by clonal deletion⁷, if they develop in an animal expressing the antigen. These studies have shown the importance and efficiency of clonal deletion of developing T cells, and have put our knowledge of tolerance induction of immature B cells on a firm basis. In the 1960s, however, it was demonstrated that tolerance could be induced in mature animals, provided that antigen was introduced in the appropriate way (intravenously, for example)⁹. Moreover, it is hard to imagine how components that are apparently expressed only in selected tissues could be delivered to the sites of lymphocyte development in quantities large enough to induce clonal deletion or anergy. Thus, clonal deletion or anergy of immature lymphocytes is unlikely to be the only mechanism contributing to tolerance.

In the case of B cells, the problem of incomplete tolerance can, in principle, be sidestepped. Without the active participation of helper T cells, B cells do not make antibody responses to many antigens, including most self-components¹⁰. Antibody responses to many bacterial components do, however, occur without

helper T cells being involved¹⁰. If there is a cross-reaction between such a bacterial antigen and a self-component that failed to cause efficient clonal deletion of those B cells recognizing it, then one might expect the production of anti-self antibodies. Indeed, this appears to occur in rheumatic fever, in which antibody produced in response to streptococcal infection cross-reacts with heart tissue, causing tissue destruction.

Goodnow *et al.* now demonstrate¹ that the immune system has a solution to this problem, at least in many cases. They took B cells expressing transgene-encoded membrane immunoglobulin specific for chicken lysozyme and transferred them from an environment lacking lysozyme into a host that expressed lysozyme as a self-component by virtue of a lysozyme transgene. When this was done, mature B cells encountered antigen without antigen-specific helper T cells. The result was that the B cells entered a state of profound unresponsiveness. Moreover, the B cells underwent a phenotypic change characterized by greatly decreased expression of one form of membrane immunoglobulin, mIgM, but continued expression of another form, mIgD. Goodnow and co-workers had previously shown⁶ that if the immunoglobulin transgenic B cells developed in a lysozyme-expressing mouse, then the B cells underwent clonal anergy, and showed the same phenotypic change described above. Surprisingly, there was no apparent difference between encountering antigen as immature B cells and encountering it as mature B cells. From previous work, one would have expected a considerable difference in the amount of antigen required to induce clonal anergy, with less needed for immature cells. If B cells with low amounts of mIgM and high amounts of mIgD have all undergone clonal anergy, either as immature B cells or as mature B cells, then this is a common fate, as Goodnow *et al.* report¹ that a

A second mechanism for immunological tolerance. Many (but not all) lymphocytes that recognize self-components are inactivated if, at an immature stage, they come into contact with antigen. Lymphocytes that mature without this happening may come into contact with self or foreign molecules through their antigen receptors, which leads to cellular proliferation and immune responses (antibody secretion, interleukin secretion and so on) provided appropriate second signals are also received. For B cells, such second signals can come from helper T cells (in the form of interleukins) or, in the case of T-cell independent antigens, probably from macrophages. For helper T cells, the antigen-presenting cell (typically a macrophage, B cell or dendritic cell) provides a 'co-stimulatory signal' which is as yet unknown. In either case, if mature lymphocytes come into contact with antigen in the absence of the second signal, the cell enters a prolonged period of unresponsiveness.



considerable fraction of normal splenic B cells have this phenotype.

Thus, antigen-induced unresponsiveness of mature B cells could be an important mechanism for decreasing autoantibody production. Previous studies have demonstrated that clonal deletion or anergy of immature B cells acts on higher affinity B cells but not on lower affinity B cells¹¹. If the same is true of clonal anergy of mature B cells, then the only B cells capable of producing autoantibodies would be those of low affinity. Most of the time such low-affinity antibodies are probably not destructive, although there may be exceptions, such as the autoantibodies produced in rheumatic fever.

Mature helper T cells can also become unresponsive on coming into contact with antigen. There are now several reports of situations where the immune system tolerates allogeneic class II MHC molecules (which are recognized by helper T cells) introduced by grafts or tissue-specific expression of a transgene^{9,12,13}. Clonal deletion of immature T cells probably cannot operate on T cells recognizing allogeneic MHC molecules unless those molecules are expressed on cells in the thymus (see ref. 9). So acceptance of these tissues probably occurs in the presence of helper T cells that can recognize the antigen^{9,12,14}. Similarly, adult Mls-1^b mice made unresponsive to Mls-1^a spleen cells were found to have normal numbers of helper T cells recognizing Mls-1^a, as detected with anti-V β 6 antibodies¹⁵. The T cells from unresponsive mice were atypical in that they did not proliferate or secrete interleukin-2 in response to stimulation with Mls-1^a spleen cells *in vitro*. Thus, it would appear that unresponsiveness is accompanied by antigen-specific helper T cells existing in an anergic state.

Why do these mature helper T cells fail to initiate graft rejection? One possibility is that the cells recognize antigen in such a way as to induce clonal anergy rather than clonal reactivity. Moreover, tolerance induction of mature helper T cells would go a long way towards explaining tolerance to self-components with restricted tissue distribution, and such a phenomenon has been observed *in vitro* with cloned helper T cells. If these T cells encounter their antigen-MHC complex in a sterile setting, for example in artificial planar membranes or on the surface of lightly fixed antigen-presenting cells, then a profound state of unresponsiveness to antigen is induced¹⁶. Interestingly, such cells proliferate vigorously on addition of the growth factor interleukin-2, raising the possibility that anergy may be a reversible state. The difference between induction of anergy and induction of a positive response, including lymphokine secretion, is the provision of a second signal (currently unknown) by the antigen-presenting cell.

One thing that remains unclear is what determines whether or not the antigen-presenting cell delivers the second signal to the helper T cell. Recently, transgenic mice were created in which expression of the I-E class II MHC molecule was restricted to the pancreas and kidney¹². The mice did not reject these tissues immunologically, but the presence of reactive helper T cells could be demonstrated on removal and analysis *in vitro*. This could be an example of clonal anergy induced in mature T cells. In agreement with that idea, I-E-expressing cells from the pancreas were able to induce unresponsiveness in helper T-cell clones, *in vitro*¹⁷. Perhaps the pancreatic cells lack the capacity to provide the second signal needed by the helper T cells to respond. This suggests that helper T cell recognition of antigen plus MHC on tissue cells may lead to unresponsiveness, whereas more specialized antigen-presenting cells (macrophages, dendritic cells or B cells) may be needed to induce activation.

More experiments will be needed to establish whether the clonal anergy seen *in vitro* with helper T-cell clones is also important *in vivo*. Experiments with transgenic T cells, following the strategy employed by Goodnow *et al.* with B cells, would be one approach. MHC alloantigen expressed on selected tissues may not be a good model for tissue-specific antigens, however, and more experiments are needed with non-MHC antigens expressed in a tissue-specific manner (as has been done with transgene-encoded SV40 T-antigen¹⁸).

It may be apparent to the reader that there is an underlying similarity between antigen-induced unresponsiveness of

cloned helper T cells and that of B cells. In both cases, there are two possibilities for the cell that encounters antigen. In the presence of a second signal coming from a helper T cell or from an antigen-presenting cell, the cell becomes activated; in the absence of the second signal, the cell becomes anergic (see figure). Antigen contact is never neutral. Everything is friend or foe, and the second signal is used to identify the foe. The idea that antigen in the absence of a second signal induces unresponsiveness was originally advanced to explain tolerance nearly 20 years ago by Bretscher and Cohen¹⁹. The experiments have finally begun to catch up with them. □

Anthony L. DeFranco is in the Department of Microbiology and Immunology and the George Williams Hooper Foundation, University of California, San Francisco, California 94143-0552, USA.

1. Goodnow, C.C. *et al.* *Nature* **342**, 385–391 (1989).
2. Owen, R.D., Davis, H.P. & Morgan, R.F. *J. Hered.* **37**, 291–297 (1946).
3. Nossal, G.J.V. *A. Rev. Immun.* **1**, 33–62 (1983).
4. Kisielow, P. *et al.* *Nature* **333**, 742–746 (1988).
5. Sha, W.C. *Nature* **336**, 73–76 (1988).
6. Goodnow, C.C. *et al.* *Nature* **334**, 676–682 (1988).
7. Nemazee, D.A. & Burki, K. *Nature* **337**, 562–566 (1989).
8. Kappler, J.W., Roehm, N. & Marrack, P. *Cell* **49**, 273–280 (1987).
9. Schwartz, R.H. *Cell* **57**, 1073–1081 (1989).
10. DeFranco, A.L. *A. Rev. Cell Biol.* **3**, 143–178 (1987).
11. Riley, R.L. & Klinman, N.R. *J. Immun.* **136**, 3147–3154 (1986).
12. Nossal, G.J.V. *Science* **245**, 147–153 (1989).
13. Mueller, D.L. *Nature* **339**, 513–514 (1989).
14. Lo, D. *et al.* *Cell* **53**, 159–168 (1988).
15. Rammensee, H.-G., Krohewski, R. & Frangoulis, B. *Nature* **339**, 541–544 (1989).
16. Mueller, D.L., Jenkins, M.K. & Schwartz, R.J. *A. Rev. Immun.* **7**, 445–480 (1989).
17. Markmann, J. *et al.* *Nature* **336**, 476–479 (1988).
18. Adams, T.E., Alpert, S. & Hanahan, D. *Nature* **325**, 223–228 (1987).
19. Bretscher, P. & Cohen, M. *Science* **169**, 1042–1049 (1970).

ASTRONOMY

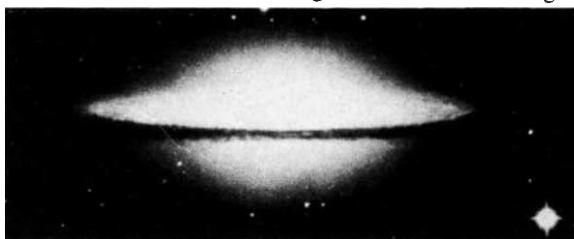
Galactic disk loses weight

M. G. Edmunds

ASTRONOMERS are so embarrassed these days by finding evidence of hidden mass almost everywhere they look, that it is a refreshing change to read a triplet of papers¹ that implies that, although the existence of dark matter in galaxies still

field, the structure of which is set by the distribution of mass. Three major components of mass can be identified. From visual and radio observation we know a great deal about the disk of stars and interstellar gas, which rotates inside a spheroid of stars (see figure). The third, invisible, component is the corona (sometimes called the dark halo) whose presence is inferred only by its gravitational influence on the rotation of the disk.

By looking out parallel to the galactic rotation axis from our peripheral position in the galactic disk, one sees a change in number density and velocity dispersion of stars as a function of distance from the plane, which can be used to investigate the gravitational field. The greatest influence in this



Sombrero galaxy with dense, bright disk and inner spheroid.

seems inevitable, it is no longer necessary to invoke its presence as a major constituent of the disk of our own Galaxy.

The dynamics of stars and gas in the Galaxy are governed by its gravitational

direction is the mass in the disk. In contrast, the variation of rotation velocity in the disk with distance from the centre (the rotation curve) is dominated by the spheroidally distributed mass of the spheroid and corona.

Kuijken and Gilmore¹ have extensively modelled the stellar density and velocity data perpendicular to the plane, to arrive at a value for the integrated surface mass density in a column through the galactic disk at the Sun. The mass in the disk determined in this way (by the dynamics of stars perpendicular to the galactic plane) is sometimes known as the Oort limit, after the pioneering investigations² of the Dutch astronomer Jan Oort. Other investigations have indicated about 50 per cent more mass detected dynamically than could be seen in stars or gas. The new work fits models to the data in a rather more general way, and also takes account of the (previously neglected) effect of the increasing influence of the spheroidal gravitational field on stars that reach high above the plane. A correction is also applied in estimating the number of stars at different heights to take account of the effects of chemical composition of stars on their brightness. Because stars that are far from the disk have, on average, a lower abundance of heavy elements than those close to the disk, neglect of this difference could lead to systematic errors in estimating distances and space densities.

The result of the investigations is a local surface density of 46 ± 9 solar masses per unit area (square kiloparsecs) of disk. This is considerably smaller than previous values. The authors give an estimate in the same units of visible matter in the disk as 35 ± 5 for stars and 13 ± 3 for gas. This totals 48 ± 6 , in excellent agreement with the dynamical mass. Thus, although some 10 units of dark matter could still be lurking in the disk without offending the error limits, the importance of this investigation is that it implies there is no need to invoke any dark matter in the disk.

The low limit on disk mass has other implications. The rotation curve from the Galactic Centre to about 50 per cent beyond our radius from the centre is reasonably well determined and can be explained only if there is a dark massive corona. Models have been constructed that try to dispense with the corona but still account for the rotation curve by putting as much mass as possible into the disk and visible spheroid — so called 'maximal disk' models³. The new local disk mass limit means that the disk itself cannot be massive enough. Although still not constrained as much as some people would like to believe, the visible spheroid probably cannot provide sufficient mass either.

The rotation curve beyond one-and-a-half times our galactocentric radius is less well measured, but certainly implies a dominant distribution of mass unlike that

of the spheroid. The reasons for the uncertainty in spheroid mass are both controversy over its shape, and also because it is difficult observationally to measure its density. The problem arises because there are very few spheroid stars in the local solar neighbourhood where they are overwhelmingly outnumbered by disk stars, although the overall spheroid could still contain a lot of stars in total because of its large volume.

The elimination of dark matter from the disk is a pleasant simplification, because we now have only to worry about the corona, a problem well-known in external spiral galaxies. But the shape and other properties of the corona remain obscure, although Kuijken and Gilmore argue that significant flattening of the corona can be

PALAEONTOLOGY

Ears and vertebrate evolution

Alec L. Panchen

IN 1987 the members of a highly successful Anglo-Danish expedition returned from east Greenland with about three-quarters of a tonne of rock matrix containing Upper Devonian fossil tetrapods (land vertebrates)¹. East Greenland is the only area in the world to have yielded tetrapods of this early date (about 360 million years old) in substantial quantities, but most of the material previously collected is attributed to the genus *Ichthyostega*². The majority of the specimens from the recent expedition appear to be of the very different genus *Acanthostega*, which was formerly known only from a handful of specimens. On page 425 of this issue Clack³, an organizing member of the expedition, describes the stapes of *Acanthostega* in a first account of the new material. She proposes that the stapes in this very early tetrapod was concerned not with the transmission of air-borne sound but with respiratory movements through a spiracle, as in some fishes. This hypothesis adds a further twist to the seemingly endless controversy over the evolution of hearing in tetrapods.

In most living tetrapods the stapes conducts air-borne sound from the tympanum (eardrum) to the inner ear. In three tetrapod groups — frogs, living reptiles and birds — the stapes is inserted in the tympanum and is the only ear ossicle. It is generally agreed that it is the homologue, at least in part, of the hyomandibula of fishes. In fishes, the hyomandibula, a part of a gill arch, serves to suspend the jaw articulation from the braincase on each side of the skull. In mammals, the middle ear contains not one ear ossicle, but a chain of three (stapes, incus and malleus), the last of which is inserted into the tympanum. Discussion of the homo-

ruled out as it would tend to reduce their dynamical estimate of disk mass below that actually visible in stars and gas. One useful additional constraint they point out is that the modelling of dynamics perpendicular to the disk effectively rules out any model that dispenses with a corona and explains the rotation curve by a modification of the law of gravitation to include an extra source of acceleration parallel to the newtonian force. □

M. G. Edmunds is in the Department of Physics, University of Wales, PO Box 913, Cardiff CF1 3TH, UK.

1. Kuijken, K. & Gilmore, G. *Mon. Not. R. astr. Soc.* **239**, 571–603, 605–649 & 651–664 (1989).

2. Oort, J. *Bull. Astr. Inst. Netherlands* **6**, 249 (1932).

3. Sellwood, J. A. & Sanders, R. H. *Mon. Not. R. astr. Soc.* **233**, 611–620 (1988).

logies of these three ossicles started well before the publication of Darwin's *Origin of Species*, when, in 1837, Reichert showed from embryological evidence that the incus and malleus were the homologues of the cartilages that form the jaw joint in all other jawed vertebrates⁴. Reichert's theory was later interpreted in evolutionary terms and is corroborated by evidence from the fossil forebears of mammals⁵. This theory is generally (but not universally²) accepted today.

It also seems certain, however, that the nearest common ancestor of mammals on one hand, and reptiles and birds on the other, did not have a tympanic ear. In those fossils believed to be close to that common ancestor (that is, the most primitive amniotes) the orientation of the stapes is like that of the fish hyomandibula from which it was derived, and there are no signs of the characteristic otic notches at the back of the skull that held the tympanum in other early tetrapods. Furthermore, there is evidence that the tympanic ear of frogs evolved separately from that of any amniote⁶. Thus a tympanic ear must have evolved convergently in three unrelated groups of tetrapods — frogs, reptiles and birds, and mammals.

Frogs, and more doubtfully other extant amphibia, are members of a larger group of primitive tetrapods, the temnospondyls. Most fossil temnospondyls have a pair of otic notches at the back of the skull roof and a lightly built stapes directed towards each notch. It is thus probable that they had tympanic ears. But the most primitive temnospondyls, notably *Greererpeton* from the Carboniferous period, had a massive stapes with a skull-bracing function⁷. Clack has shown that another Carboniferous tetrapod, *Pholidrpeton*, which is thought to be related

to the amniote line, has a strikingly similar stapes⁸. Both *Greerpeton* and *Pholidroperon* lack otic notches.

Clack now shows³ that the stapes of the much earlier *Acanthostega* is of the same type. Yet *Acanthostega* also appears to have otic notches. She suggests, however, that these did not hold tympana, but that each formed the margin of an open spiracle. She further suggests that the stapes was concerned with pumping movements to draw in either air or water through the spiracle for respiration. In fish where the spiracle remains open, notably in skates and rays, the hyomandibula helps to pump water into the buccal cavity; the water is then expelled through the other gill slits. Clack suggests that the stapes of *Acanthostega*, the homologue of the hyomandibula, also had this respiratory function, but probably to ventilate the lungs.

If Clack and her colleagues had not recovered stapes associated with skull material, one might have predicted that the function of *Acanthostega*'s stapes was to conduct sound from a tympanum held in the otic notch. Given that three unre-

lated tetrapod groups are known to have this arrangement, it would have been reasonable to suppose that this was the primitive tetrapod condition and thus would be found in *Acanthostega*. The lack of otic notches, and the massive stapes found in *Greerpeton* and *Pholidroperon*, could then have been seen as secondary features. The discovery in *Acanthostega* of both massive stapes and otic notches thus forces a re-evaluation of the primitive tetrapod condition that will inevitably affect our ideas about tetrapod phylogeny. □

Alec L. Panchen is in the Department of Biology, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, UK.

1. Panchen, A.L. *Nature* **333**, 704 (1988).
2. Jarvik, E. *Basic Structure and Evolution of Vertebrates* (Academic, London, 1980).
3. Clack, J.A. *Nature* **342**, 425–427 (1989).
4. Goodrich, E.S. *Studies on the Structure and Development of Vertebrates* (Macmillan, London, 1930).
5. Kemp, T.S. *Mammal-like Reptiles and the Origin of Mammals* (Academic, London, 1982).
6. Lombard, R.E. & Bolt, J.R. *Biol. J. Linn. Soc.* **11**, 19–76 (1979).
7. Smithson, T.R. *Zool. J. Linn. Soc.* **76**, 29–90 (1982).
8. Clack, J.A. *Zool. J. Linn. Soc.* **79**, 121–148 (1983).

AMAZON DEFORESTATION

Predicting climate effects

Robert E. Dickinson

FEARS that global and local meteorological balances might be upset by the deforestation of the Amazon have made the destruction of the rainforest a matter of wide public concern. Elsewhere in this issue (*Nature* **342**, 411–413; 1989), J. Lean and D. Warrilow describe detailed computer simulations that reveal the impact that might be expected if large tracts of rainforest are replaced by pasture. Most strikingly, they find that rainfall in the region could be dramatically reduced.

The role of tropical forests in the global climate system has long been a subject of speculation. Brazilian studies of the hydrological balance in the Amazon have done much to foster the debate over the past decade. Joint UK–Brazilian micro-meteorological investigations are now giving us a much clearer view of the balances of energy and water in and above the forest canopy. But it is only with detailed climate models, such as those used by Lean and Warrilow and also by C. Nobre, J. Shukla and P. Sellers (submitted to *Science*), including representations of the role of vegetation in the water and surface-energy balances, that we can make quantitative estimates of the regional and global changes that may follow extensive deforestation.

As well as predicting a 20 per cent decrease in rainfall over the Amazon region, Lean and Warrilow suggest that

the local climate will become warmer and that evapo-transpiration will be reduced. My earlier study with A. Henderson-Sellers and M.F. Wilson (see *Q. Jl R. met. Soc.* **114**, 439–462; 1988), using the National Center for Atmospheric Research community climate model (CCM), also indicated that warming and reduction of evapo-transpiration on a similar scale would follow deforestation, but the model hydrological cycle was too 'noisy' to establish any clear indication of

changes in rainfall.

Regional rainfall in the Amazon is highly variable, in both space and time, owing to variations in tropical ocean temperatures. Until climate models can reproduce this variability, projections of changes induced by altered land use will be greeted with scepticism. The hydrological cycle is regarded by many as the weakest component in climate models. The main difference between our earlier study and Lean and Warrilow's is in the numerical treatment of the spatial dimensions, their approach avoiding hydrological difficulties associated with ours. The link the new study seems to have established between rainfall and tropical forests is one of the strongest of its kind yet modelled.

The suggestion of this link increases the urgency with which related questions must be addressed. How do the rainfall anomalies affect weather patterns in the Northern Hemisphere? By analogy with the El Niño–Southern Oscillation phenomena, large-wavelength, transverse 'Rossby' waves in the atmosphere might propagate long-range effects. Will the accompanying changes in ocean temperatures amplify or mitigate the predicted effects? It will be important to find out whether climate changes due to deforestation will act through the forest ecosystem to accelerate the current decline. It may be that the tropical forests are sufficiently stable that they will grow back once human intervention ceases.

The interest in the rainforest reminds us of the other effect of vegetation on climate. Vegetation is a reservoir for carbon that annually exchanges with the atmosphere an order of magnitude more CO₂ than is added by human activity. The loss of the tropical forests is recognized as adding significantly to the increases in atmospheric CO₂. As the political will grows to prevent such increases, it is

100 years ago

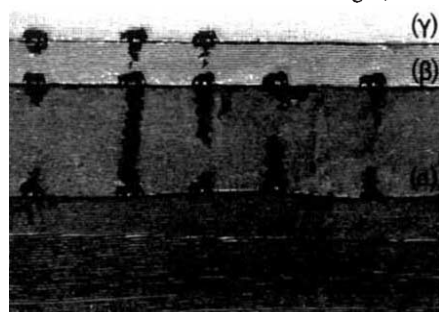
Mirage in the South American Pampas

I was staying in the Pampas of the Argentine Republic, near Melincue, a small town of the Province of Santa Fe, from September 1888 to March 1889. During my stay I had the opportunity of observing certain mirage phenomena. To illustrate my observations I had drawn eight diagrams; but, for the purpose of insertion in *Nature*, I have been obliged to reduce these. Hence I fear that my descriptions may not be as clear as I should wish.

The Winter Mirage.

I saw this mirage several times, always about sunrise and after a frost. There would be, for example, just below, or on the edge of the line (α), a cow. This I will call the "first cow," or the "original cow." Just below or on the line (β), vertically above the first cow, and, like it, erect, would be a second cow, a repetition of the first. And often, above this again, below or on the line (γ), would be a third cow, also erect.

Sometimes there were confused images hanging from the second cow and joining other confused images piled on the first cow; sometimes the first cow was clear of images, while



they hung down from the second cow; sometimes the second cow was near, and there were images piled on the first.

The mirage lasted until about an hour and a quarter after sunrise.

W. LARDEN

From *Nature* **XLI**, 69–71; November 1889.

worth noting that the cost of techniques for reducing CO₂ emissions, up to many hundreds of dollars per tonne of carbon removed, probably outweighs any economic benefit gained from forest destruction. It is easy to imagine that the path with least initial cost to reduce CO₂ inputs to the atmosphere would be in the form of payments to countries with tropical forests for a reduction in cutting and burning.

The many political, social and eco-

logical factors behind tropical deforestation make this a complex issue. But although studies of the climatic impact of deforestation will help climate modellers understand the important features in the weather system, it is questionable that they will help policy makers in tropical countries tackle their ecological problems. □

Robert E. Dickinson is at the National Center for Atmospheric Research, PO Box 3000, Boulder, Colorado 80307, USA.

MID-OCEAN RIDGES

Mixing of basalt magmas

Emily M. Klein

THE pressure at which melting and melt segregation occur to produce mid-ocean ridge basalt (MORB) magmas is at the heart of a long-running controversy. This is due to the fact that from the pressure one can constrain physical characteristics of divergent plate boundaries, such as the temperature structure beneath the ridge axis or the width of the melting zone, and chemical characteristics of the primary magmas in equilibrium with the mantle source. Some argue¹ that primitive MORBs represent nearly primary magmas in equilibrium with the mantle at pressures as low as 7 kbar, corresponding to a relatively shallow depth, 21 km, beneath the ridge axis. Others suggest² that extensive crystallization *en route* to the sea floor has substantially modified the primary melt that was generated at pressures of 15–30 kbar (45–90 km). Although the two schools may seem irreconcilable, Salters and Hart, on page 420 of this issue³, contribute to a growing body of evidence suggesting that advocates of both extremes may, in fact, be correct, united under a more realistic model for the melting process beneath ocean ridges.

In their study, Salters and Hart examine Lu–Hf and Sm–Nd isotope and trace-element systematics in MORB and ocean-island basalts. The authors identify a discrepancy between the measured parent-to-daughter trace-element ratios of Sm/Nd and Lu/Hf, on the one hand, and their isotope ratios ¹⁴³Nd/¹⁴⁴Nd and ¹⁷⁶Hf/¹⁷⁷Hf, on the other. The observed isotope ratios are ‘unsupported’ in the sense that the high ratios imply parent-to-daughter trace-element ratios higher than observed. For the Lu–Hf system, the authors term this discrepancy the Hf-paradox and explore the possibility that the decoupling between trace-element and isotope ratios occurs during the melting process itself.

Specifically, if garnet is a residual phase during the melting process beneath the ocean ridge, the greater chemical affinity of garnet for Lu than for Hf will lead to low Lu/Hf ratios in the melts produced within the garnet stability field. The

stability of garnet as a residual phase in turn implies that a significant portion of the melting process occurs at pressures greater than about 20 kbar, where garnet replaces spinel as the stable aluminous mantle phase.

Salters and Hart adopt a more realistic melting model that recognizes that if melting occurs by adiabatic decompression, it must be polybaric, with melts generated (and segregated) over a range of pressures beginning deep within the garnet stability field and extending to lower pressures within the spinel stability field. The basaltic magma erupted at the ridge axis thus represents an average composition generated by the pooling of melts from throughout the polybaric melting column. From this perspective, advocates of high- or low-pressure melting are both correct in

the sense that melting occurs over a range of pressures. It is important to recognize, however, that for a pooled melt composition, the concept of a single pressure of equilibration during melting, as implied in the earlier studies, has no meaning.

Studies of the major-element compositions of MORB also support the concept of a polybaric melting column and the conclusion that for the majority of MORB magmas, melting begins at high pressures (20–40 kbar). Examining the range of major-element compositions observed among regional averages of MORB chemistry, Langmuir and I noted⁴ that parameters indicative of the extent of melting, such as sodium content or the ratio of calcium to aluminium, correlate with parameters indicative of the pressure of melting, such as iron or silica.

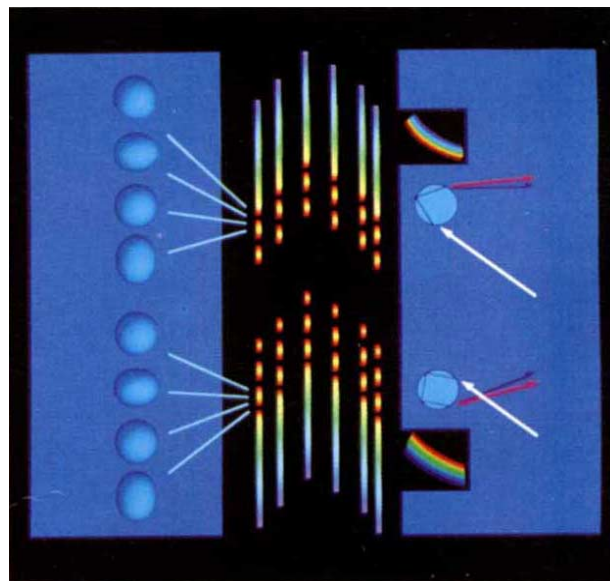
As with the Lu–Hf systematics, this relationship can most easily be understood in terms of the pooling of melts from throughout polybaric melting columns; the height of each column is governed by the solidus pressure and therefore the mantle temperature beneath the ridge axis for a given region. The observed correlation between the extent of melting and pressure of melting results from the fact that the deeper the intersection of the mantle solidus, the greater the proportion of melt generated at high pressure and the greater the total amount of melting achieved. Indeed, Salters and Hart note the complementary nature of the major-element, trace-element and isotope data and observe that their indicator of the pressure of melting, $\delta(\text{Lu/Hf})$, correlates

Rainbows reveal good vibrations

How big is a raindrop?

One way to find out is to use radar. Air pressure flattens falling drops, more so for larger drops, the oblateness being detectable with polarized radar beams. But, as Kenneth Beard and colleagues point out on page 408, oscillations of the raindrops, by averaging effects can make them seem smaller. This effect becomes noticeable for drops 1–1.5 mm in diameter, possibly owing to a resonance between the natural frequency of drops this size (a few hundred hertz) and the

frequency with which air eddies in the wake are shed. To test this hypothesis, Beard *et al.* observed the changing rainbow reflections of white light scattered by each drop as it fell. Lamp and detector are arranged to favour red light reflections. As the drop oscillates, either yellow or infrared (dark) light is preferentially reflected, leaving a characteristic pattern on streak photographs of the falling drops, as depicted above (top, primary rainbow; bottom, secondary). □



with major-element indicators of the extent of melting, such as sodium.

The idea that the generation of MORB involves the pooling of melts is also supported indirectly by a study of Sr, Nd and Pb isotopes in basalts recovered from near-ridge seamounts and their adjacent ridge segments, reported by Castillo and Batiza in *Nature* last week⁵. Previous studies of the East Pacific Rise axis and adjacent seamount isotope systematics have shown that, although the two environments overlap in isotope composition, the isotope diversity of the seamount lavas is considerably greater than that of the ridge-axis lavas⁶. The proximity of the ridge and seamounts suggests a similar mantle source, composed of dispersed, isotopically heterogeneous mantle material.

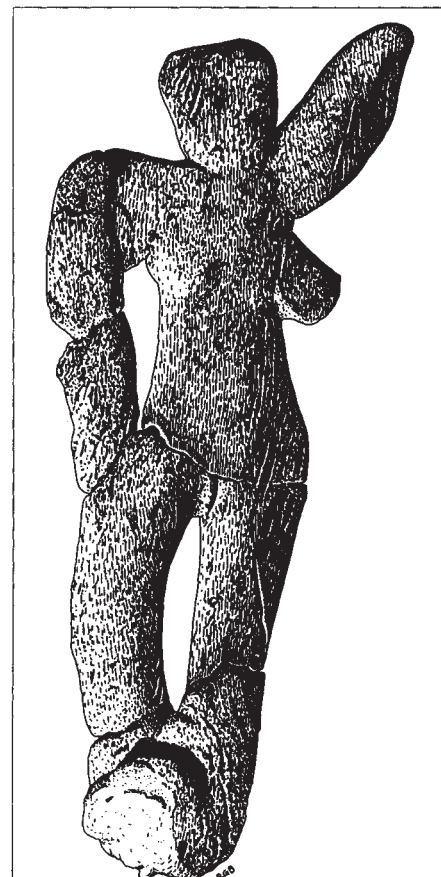
The greater isotopic diversity of the seamounts may therefore result from the sampling of isotopically distinct reservoirs and eruption of the distinct magma compositions, while the greater homogeneity of the ridge-axis magmas may result from the pooling of these isotopically diverse melts before eruption. This interpretation is easy to fathom for the fast-spreading Pacific ridges, where robust magma supply and the presence of a subaxial magma chamber⁷ can be expected to foster homogenization of diverse melt compositions.

Castillo and Batiza show that the relationship observed in the Pacific also holds for seamounts and the adjacent segment

of the mid-Atlantic ridge in the South Atlantic. This is despite the fact that the spreading rate of the southern mid-Atlantic ridge is one-quarter that of the East Pacific Rise, that magma supply is presumably less voluminous, and that evidence of a steady-state magma chamber to facilitate homogenization has yet to be documented. Nevertheless, the isotope study of Castillo and Batiza suggests that significant mixing of isotopically distinct melts occurs even beneath the mid-Atlantic ridge, attesting to the homogenization of melts even at slow spreading rates. Although it is not clear where this homogenization process is taking place, whether in magma conduits or perhaps in small, ephemeral magma chambers, there is now a substantial body of evidence from isotopes, trace elements and major elements that MORB-magma genesis involves the pooling of melts from a heterogeneous mantle over a wide range of pressures. □

Emily M. Klein is in the Department of Geology, Duke University, Durham, North Carolina 27708, USA.

1. Presnall, D.C. & Hoover, J.D. *Contrib. Mineral. Petrol.* **87**, 170–178 (1984).
2. O'Hara, M.J. *Nature* **220**, 683–686 (1968).
3. Salters, V.J.M. & Hart, S.R. *Nature* **342**, 420–422 (1989).
4. Klein, E.M. & Langmuir, C.H. *J. geophys. Res.* **92**, 8089–8115 (1987).
5. Castillo, P. & Batiza, R. *Nature* **342**, 262–265 (1989).
6. Zindler, A., Staudigel, H. and Batiza, R. *Earth planet. Sci. Lett.* **70**, 175–195 (1984).
7. Detrick, R.S. *et al.* *Nature* **326**, 35–41 (1987).



The 'Dancing Venus of Galgenberg', which is some 30,000 years old. The sculpture is 7.2 cm long and is carved from green serpentine stone. (Drawing by R. Bednarik, from a photograph.)

ARCHAEOLOGY

Age and the female form

Paul G. Bahn

ARCHAEOLOGISTS in Austria have discovered one of the world's earliest known sculptures, in the form of the oldest dated female figurine¹. The find is remarkable not merely for its antiquity but also for the artistic sophistication it displays.

Since 1941 the loess sediments of the Galgenberg site, near Krems, have been known to contain bones of Ice Age mammals (including reindeer and mammoth) together with abundant charcoal and flint flakes. New excavations in a nearby vineyard, directed by Christine Neugebauer-Maresch for the Österreichisches Bundesdenkmalamt, led to the discovery in September 1988 of the fragments of the female figurine. Charcoal samples from the same layer have produced radiocarbon dates of 29,200 to 31,190 years before present (BP).

The 'Dancing Venus of Galgenberg', as the sculpture has been dubbed, is 7.2 cm long, and carved in green serpentine. Unlike the well-known 'Venus figurines' of the Ice Age, such as that from Willendorf in Austria (found almost exactly 80 years earlier to the day), the Galgenberg

specimen is flattish rather than carved in the round; its shape was probably predetermined by the stone used, because serpentine often occurs in slabs. The figure has not been polished, and its surface shows the marks of tools and erosion.

The right arm and the legs are supported at both ends, while the left arm appears to be folded back at the elbow. Body weight is supported primarily on the left leg, while the right leg rests on a slightly higher support. The right hand is placed on the hip. This pose causes the left breast to be depicted almost in profile, while the right is in very low relief because of the stone's flatness. The vulva is marked (a feature rare in Ice Age depictions of females²), and in further contrast to many of the better known 'Venus figurines' there is no hint of obesity or of an emphasis on breasts and buttocks. The figure is more or less anatomically accurate, except for the thickened limbs which are probably as thin as the artist dared carve them without making the sculpture unduly fragile.

There are two important points about

the Galgenberg figurine. First, it is at least 5,000 years older than the 'Gravettian period' to which most 'Venus figurines' are usually, often unjustifiably, assigned; most of the Western European specimens have no archaeological context whatsoever, and are merely assumed to be Gravettian³, whereas dated specimens in Eastern Europe in fact span a period from about 23,000 to 12,000 years BP⁴.

Second, the Galgenberg carving is the result of considerable technological skill — the stone is rather delicate and brittle, and the head, left arm and breast could all easily have broken off. In addition, the two openings (under one arm and between the legs) would have required a delicate boring operation; the later 'Venus figurines' are normally static, solid symmetrical figures with no perforations and no protruding limbs⁵.

The sophistication revealed in both technology and composition can only be the end-result of a long tradition in carving (perhaps of perishable materials). The same conclusion had already been reached from study of the contemporary or even older ivory carvings of animals and humans from West Germany (Vogelherd, Geissenklösterle, Hohlenstein-Stadel)⁶ and perhaps also the ivory human figurines from Brassempouy in France⁷.

These objects, the oldest figurative carvings known at present, simply cannot represent the 'first art', which must therefore have appeared long before.

In this respect, the Acheulian female 'figurine' from Berekhat Ram, Israel⁸ may provide a crucial clue. Dating to between 230,000 and 800,000 years BP, it is a small scoria pebble whose shape, naturally resembling a 'Venus figure', may have been modified by several grooves. Whether or not it was modified, its very presence at Berekhat Ram implies that the occupants of the site recognized its iconic properties, and therefore that an

interest in the female form is one of the most enduring features of mankind. □

Paul G. Bahn, 428 Anlaby Road, Hull HU3 6QP, UK, is a freelance writer on archaeology.

1. Bednarik, R.G. *Rock Art Res.* **6**, 118–125 (1989).
2. Bahn, P.G. *Rock Art Res.* **3**, 99–105 (1986).
3. Bahn, P.G. & Vertut, J. *Images of the Ice Age 85* (Windward, Leicester/Facts on File, New York, 1988).
4. Soffer, O. in *The Pleistocene Old World* (ed. Soffer, O.) 335 (Plenum, New York, 1987).
5. Delporte, H. *L'Image de la Femme dans l'Art Préhistorique* (Picard, Paris, 1979).
6. Marshack, A. in *L'Homme de Néandertal: 5, La Pensée* (ed. M. Otte) 57–91 (ERAUL, Liège, 1988).
7. Bahn, P.G. & Vertut, J. *Images of the Ice Age 140* (Windward, Leicester/Facts on File, New York, 1988).
8. Goren-Inbar, N. *Mi' Tefukat Ha'Even* **19**, 7–12 (1986).

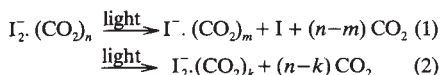
CHEMICAL PHYSICS

Size-selected solvent cages

Tony Stace

AN isolated diatomic molecule photo-excited to an electronic state above the dissociation limit will swiftly break up into free atoms (Fig. 1). In a liquid or dense gas, however, the relative kinetic energy of the excited atom pair is often lost to a surrounding solvent cage before it can separate. This is the cage effect, first proposed by Franck and Rabinowitsch¹ to explain the reduced photochemical yield, ϕ , of free radicals (atoms or molecular fragments with an unpaired electron) in solution. Attempts in liquid-phase experiments² to reveal details of the process at a molecular level are impeded by the bulk physical properties of the liquids. In an effort to resolve this problem, Lineberger and colleagues³ have studied the dynamics of atom recombination (caging) in mass-selected cluster ions.

Small clusters comprising a chromophore, I_2^- (which absorbs light at a wavelength of 720 nm) and 1–20 carbon dioxide molecules in the form $I_2^-(CO_2)_n$, generated in a vacuum at the entrance to a mass spectrometer, are mass-selected and excited by a pulsed dye laser. Depending on the value of n , this induces either of the following photofragmentation steps:



In reaction (1), I^- has photodissociated in much the same way that it might as an isolated gas-phase species. But as n increases, there is a tendency for reaction (2) to dominate: under those circumstances, the CO_2 molecules are operating as an effective cage. They absorb the kinetic energy of the separating fragments, and dissipate it by being evaporated from the cluster⁴. Provided the energy of I_2^- falls below the bond energy for the ground state (Fig. 1), the molecular ion is stabilized and cannot dissociate. Plotted in Fig. 2 is the product fraction resulting from reaction (2) as a function of cluster size;

in effect, the graph is a plot of $1 - \phi$. As might be expected, increasing the size of the cluster increases the probability of caging. In previous experiments, Lineberger's group⁵ showed that, for a given size of cluster, the photofragment yield, ϕ , increases with photon energy.

Since the early work by Franck and Rabinowitsch¹, there have been many attempts to characterize the cage effect in terms of both timescale and molecular or solvent parameters. These efforts were limited by the fact that it was not possible to monitor the influence of a single solvent variable: changing the solvent molecular mass invariably leads to a change in solvent size and/or viscosity. But about 15 years ago Troe and co-workers² used simple solvents, such as condensed inert gases, to allow a more systematic variation of parameters; thus a change from liquid argon to liquid xenon gives a considerable difference in mass, but not in viscosity. Also, the photochemical yield could be monitored as a function of gas pressure. In this way, Troe and colleagues spanned the complete range of behaviour, from low pressures for which $\phi \approx 1$, through to very high pressures (10,000 atmospheres) where little or no photodissociation was observed.

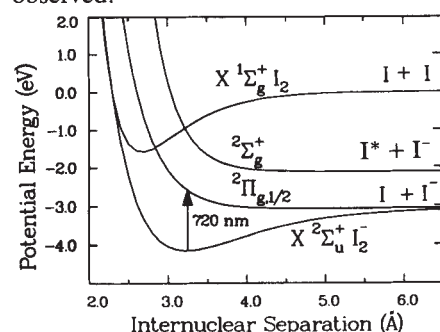


FIG. 1 Approximate potential-energy curves for I^- and I_2 . In the clusters the $2\Pi_{g,3/2} \leftarrow 2\Sigma_u^+$ transition between the bound ground state and a low dissociating state is excited with 720 nm laser radiation.

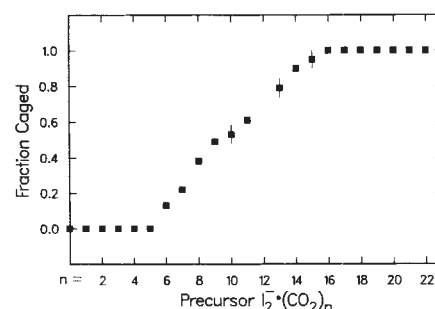


FIG. 2 Fraction of products resulting from reaction (2) as a function of cluster size, n , following the photoexcitation of $I_2^-(CO_2)_n$ with laser radiation of wavelength 720 nm. (Lineberger *et al.*, unpublished results.)

As might be expected, dense Xe, N_2 and CO_2 act as relatively efficient cages, whereas He and Ne do not. At the same time, picosecond (10^{-12} s) time-resolved measurements of the photodissociation of I_2 in liquid by Chuang *et al.*⁶ greatly elucidated the dynamics of the caging process. In particular, they led to the development of models for caging based on a diffusion-controlled recombination of geminate (original) pairs.

The current experiments by Lineberger and co-workers can be seen as a natural step forward from both these techniques. The group seeks to take an individual localized environment, as seen by the excited molecule or ion, and establish how effective it is at promoting a cage event. The latest experiments are time-resolved, and yield a measure of how quickly ground-state I_2^- in reaction (2), recovers its original electronic and vibrational state. The measured time, just 10 ± 5 ps for $I_2^-(CO_2)_{16}$ (about 500 CO_2 vibrational periods), is a convoluted response derived from both the experimental situation and instrumental parameters.

On a picosecond timescale, there are several factors associated with the physics of photodissociation and recombination that could influence events: curve-crossing (electronic-to-translational energy exchange) from the excited electronic state to the final ground-state; vibrational relaxation of the ground state (the electronic absorption profile of I_2^- will change according to its vibrational energy content); and the possibility of recombination into one of several low-lying electronic energy states (a complete description of the electronic states of I_2^- is more complex than that given in Fig. 1). Taking into consideration the results depicted in Fig. 2, one might conclude that it requires 12 or more molecules to establish an effective cage, and that under those circumstances, recombination of the iodine atoms is very rapid.

In a liquid, a small chromophore such as I_2^- could have between two and ten immediate neighbours, depending on the size of the solvent molecules. Therefore, the cluster results support the view that

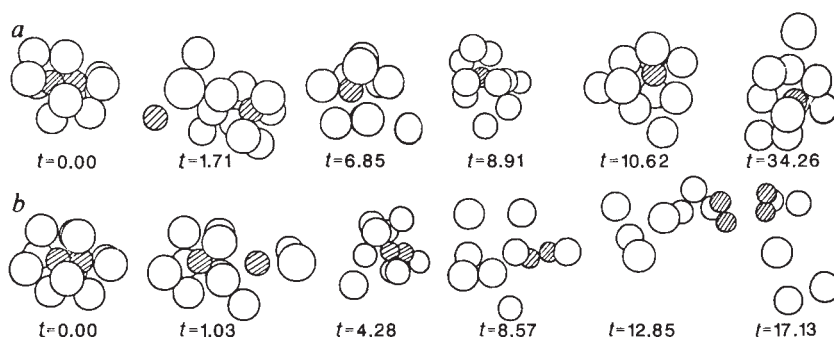


FIG. 3 Results of a computer simulation⁷ of events which immediately follow the photoexcitation of $\text{Br}_2 \cdot (\text{CO}_2)_{11}$; the bromine atoms are shaded. *a*, The Br_2 photodissociates; *b*, the cage event leads to the cluster losing most of its carbon dioxide molecules. (Time, *t*, in picoseconds.)

liquid-phase caging⁶ is a fast, localized phenomenon, and that most events are probably geminate processes. But a direct extrapolation from cluster to liquid is not necessarily valid: energy flow from a cluster is an irreversible process that has no liquid-phase analogue¹.

Parallel to the experimental progress has been the development of theoretical techniques designed to simulate the caging process at a molecular level². Many of these have evolved from advances in the computer simulation of simple liquids. As one might expect, it is easier to simulate liquid argon than it is liquid benzene or hexane, so that the experiments of Troe and co-workers² have provided an ideal source of experimental data. One obstacle associated with the computer simulation of simple liquids remains, namely boundary conditions. The computer, both by virtue of its finite memory and lifetime of the operator, can handle only a limited number of particles — typically 250–1,000 for a physical timespan of about 10^{-10} s. To achieve macroscopic proportions, an 'infinite' three-dimensional lattice is created by defining boundaries that are translated images of the particles under consideration.

The beauty of Lineberger's experiments is that, at least for the simulation of the cage effect, such artificial barriers are no longer necessary. The dynamics of the caging process in a simple cluster, $\text{I}_2 \cdot (\text{CO}_2)_n$ or $\text{Br}_2 \cdot (\text{CO}_2)_n$, can be followed using numerical techniques to solve Hamilton's classical equations of motion for the complete system. Figure 3 shows two examples where a computer has been used⁷ to simulate the sequence of steps which immediately follow the photoexcitation of $\text{Br}_2 \cdot (\text{CO}_2)_{11}$ (for the purposes of the present discussion, the behaviour of $\text{I}_2 \cdot (\text{CO}_2)_n$ and $\text{Br}_2 \cdot (\text{CO}_2)_n$ is identical). At time $t = 0$, the Br_2 chromophore absorbs a photon which increases the total energy of the cluster by approximately 1.7 eV. The temporal evolution of the cluster is then followed for up to 34 ps, during which time either of reaction steps (1) and (2) can occur. In the simulation depicted in Fig. 3*a* an ion photodissociates, and a Br atom can be seen leaving the cluster after

approximately 1.7 ps. In contrast, Fig. 3*b* shows a caging event where the excess energy in the chromophore has been successfully removed, but at the expense of almost all the CO_2 molecules contained in the cluster. The timescale for the latter, 17 ps, is in reasonable agreement with that determined by Lineberger and co-workers.

Computer techniques are available that allow simple molecules, such as N_2 and CO_2 , to be treated as structured, rigid, rotating molecules in simulations of this

DNA FINGERPRINTS

Victims or perpetrators of DNA turnover?

Gabriel A. Dover

As revolutions go, the discovery of restless minisatellite DNA families (the basis of DNA fingerprinting in humans)¹ is here to stay; and as with all good revolutions it has incited the search for other unruly sequences. As the evidence emerges, it seems that human minisatellite DNA is at the tip of an iceberg of many types of more passively fluctuating repetitive sequences.

It was no accident that Jeffreys called regions of hypervariable length fluctuations 'minisatellite' DNA, because of their similarity in organization to the large blocks of maxisatellite DNAs. The ease of isolation and sequencing of maxisatellites quickly gave investigators a panoramic view of a landscape seemingly populated by every conceivable type of satellite DNA, in terms of repeat-unit length (2–2,000 base pairs (bp)), sequence composition, copy number and chromosomal distribution. Are minisatellites as diverse and bewildering as the maxisatellites and multigene families?

The first analysis of human minisatellite repeating units exposed a conserved 'core' sequence with similarities to the *chi* recombination signal in *Escherichia coli*¹ and probably serving as a hotspot for meiotic unequal crossingover². Is this sequence an obligatory constituent within all DNA families showing hypervariability

type³. Although such a step might seem attractive, it may not help clarify caging. An important relaxation mechanism must be the exchange of energy from the vibrational modes of the excited chromophore to the vibrational modes of the molecules in the solvent cage. Treating the cage molecules as rigid rotors excludes such an exchange. But computer simulation techniques are a long way from being able to treat collections of vibrating molecules. To some extent, the ball is still in the experimentalist's court; theorists now need results for I_2 , Ar_n or $\text{I}_2 \cdot \text{Kr}_n$. □

Tony Stace is in the School of Chemistry and Molecular Sciences, University of Sussex, Falmer BN1 9QJ, UK.

1. Franck, J. & Rabinowitch, E. *Trans. Faraday Soc.* **30**, 120–131 (1934).
2. Schroeder, J. & Troe, J. A. *Rev. Phys. Chem.* **38**, 163–190 (1987).
3. Ray, D., Levinger, N.E., Papanikolas, J.M. & Lineberger, W.C. *J. Chem. Phys.* **91**, 6533–6534 (1989).
4. Stace, T. *Nature* **327**, 186–187 (1987).
5. Alexander, M.L., Levinger, N.E., Johnson, M.A., Ray, D. & Lineberger, W.C. *J. Chem. Phys.* **88**, 6200–6210 (1988).
6. Chuang, T.J., Hoffman, G.W. & Eisinger, K.B. *Chem. Phys. Lett.* **25**, 201–205 (1974).
7. Perera, L. & Amar, F.G. *J. Chem. Phys.* **90**, 7354–7368 (1989).
8. Stace, A.J. *J. Chem. Soc. Faraday Trans. 2* **75**, 2105–2110 (1981).

in copy number? Or are there many (possibly thousands) of repeated DNA families wildly fluctuating as a consequence of unequal crossingover and alternative DNA turnover mechanisms such as slippage, rolling circles, and those responsible for telomeric growth, which seemingly do not require *chi*-like sequences? With a nod to the 'selfish DNA' debate of the innocent early 1980s, we can ask if hypervariable minisatellites are constitutionally selfish or blissfully ignorant?

From comparisons of core sequences of the several minisatellite DNAs listed in the table (see ref. 3 and references therein) it could be that the conserved core sequence (GGCAGGAXG) (dashed underline in the table) found by Jeffreys is not widespread, and, given the G-rich composition of the majority of sequences being compared, the statistical meaning of the one remaining common sequence G-----GG (underlined) is not entirely convincing. Not only is this relic similarity even further removed from the core sequence, but it also places doubt on its functional relationship (if any) to the *chi* recombination signal (GCTGGTGG). Moreover, other AT-rich minisatellites, derived from the apolipoprotein B gene and the collagen type II gene, do not share any core homology with G-----GG or *chi*.

In similar vein, hypervariable length variation within exons can be due both to repeats containing *chi*-like sequences (as in the silk-moth chorion genes¹) or repeats not containing *chi*-like sequences (as in the human proline-rich protein genes²) (see table). The short length of the repeat (9 bp) in the former indicates slippage,

like ACNGGN and TCAGGC repeats from within the *per* gene of *D. melanogaster* and the synthetic (CAC)_n have also been used successfully as fingerprint probes, indicating that simple sequences are a rich untapped source of fluctuating length variants.

It is becoming

increasingly clear that eukaryotic nuclear genomes (including extensive regions of exons and control sequences) are largely non-random and repetitive on both a fine-grained and a larger scale³. They are played upon by half a dozen sequence-dependent and sequence-independent mechanisms of turnover. To misquote Tolstoy: all families have their ups and downs, but each suffers in its own way.

Furthermore, all the mechanisms of amplification and turnover frequently operate one on top of another. For example, slippage (and micro gene conversions) may often be churning around within human minisatellite arrays at some loci during spermiogenesis⁴, and in arrays of the *Drosophila* ribosomal RNA genes⁵, generating complex patterns which are uninterpretable using simplistic models of selection or drift which do not incorporate DNA turnover. There

are probably few long tracts of randomly organized single-copy DNA marching to the beat of a 'molecular clock'⁶. Minisatellites, whether victims or perpetrators of turnover, are seemingly the tips of icebergs of hundreds of heterogeneous families, tightly packed together throughout most of the genome. Human genome sequencers beware — the going will be tough. □

Gabriel A. Dover is in the Department of Genetics, University of Cambridge, Cambridge CB2 3EH, UK.

Molecular knitting

MANY monomers, like styrene, can polymerize ionically. The incomplete polymer chain terminates in an electric charge; a monomer molecule reacts with it, the charge jumps to the new end, and the action is repeated. So the growing head of the chain carries a charge along with it.

Now a charge moving in a magnetic field travels in a circle. So, says Daedalus, a polymer chain extending in such a field will grow in a circle. The reaction should give that elusive rarity, a closed-loop ring polymer. Random brownian displacements may prevent perfect closure on the first circuit, but sooner or later the growing, winding snake will bite his own tail.

In a dilute solution, individual ring molecules will result. But in strong solutions or pure liquid monomer, the crowded molecular loops will inevitably grow through each other. The result will be a unique 'molecular chain mail' — a three-dimensional lattice of interwoven loops.

Molecular chain mail will have intriguing properties. Pulled in one direction, it will shrink strongly in another, like a knitted pullover. It will not melt and will have a fantastic breaking strain, for it can fail only by the shearing of its chemical bonds. Made in a weak magnetic field, it will be soft and flexible, with large, floppy loop molecules; a strong field will give a tough solid with small, tight molecular links. It could even be made totally rigid by 'vulcanizing' it with a few cross links between the loops. Even so, it would not be brittle; for excessive loads would merely shear the cross links. The material would then 'work soften' locally, and yield slightly to relieve the concentration of stress.

More interesting still, molecular chain mail will have strongly nonlinear elasticity. It will yield readily to a small force, but as the molecular loops distort and pull taut, it will increasingly resist further extension. This is the subtle characteristic which gives biological tissues their remarkably high tenacity and freedom from catastrophic failure. It saves arteries, for example, from inflating locally under excessive pressure, and stops a bat's wing from tearing across if it catches on a thorn.

At last synthetic chemistry will be able to rival leather, catgut, sinew and other prized biological materials. Unbelievably pliable and strong, Daedalus's molecular chain-mail polymers will invade the market with a thousand uses, from shoes, handbags and coats to tubing, violin strings and balloon fabrics. And surgeons may welcome them as the ideal artificial material for soft-tissue repairs, fully compatible with the mechanical properties of their living surroundings. David Jones

Some representative 'core' and simple-sequences from hypervariable repeats (see also ref. 3)

G-rich core sequences

E. coli

Chi

GCTGGTGG

Human

Myoglobin

GGAGGTGGCAGGAAC

consensus 'core'

GGNNGTGGCAGGAXG

(using myoglobin probe)

Bird

(willow warbler)

minisatellite — 1

minisatellite — 2

(using myoglobin probe)

AGGGAAGGGCTC

GGGACAGGGGAC

Human

zeta globin

GGTTGTGAGTGGGGCACA

insulin

ACAGGGGTGTGGGG

Alpha globin

GAGCGACACGGGGG

Harvey-ras

GGGGAGTGTGGCGTCCC

oncogene

Phage

M13

GAGGGTGGNGNTCT

A-rich core sequences

Human

Apolipoprotein B

TTTATAATTAATATTTT

Collagen type II

CAATATAGAAATAATA

Simple sequences

probes used to

isolate

minisatellites

synthetic

human

Drosophila

whale

Drosophila per

locus

(CAC)_n

(GT)_n

(CAG)_n

(CT)_n

(ACNGGN)_n

and (TCAGGC)_n

Hypervariable exonic repeats

Chi-like?

Silk-moth chorion genes

TGYGGXGGX

(X = A, T, C)

(Y = T, C)

Not *chi*-like?

Human proline-rich

protein genes

..AGGAAAGCCACAGGACCACCCC..

whereas the longer repeat (48–63 bp) in the latter indicates unequal crossingover, as mechanisms of turnover.

The widespread occurrence and high sequence heterogeneity of slippage-generated simple sequences means that a great proportion of eukaryotic genomes have the potential to yield hypervariable arrays that are unique to one or other single locus, a situation of some importance in the application of DNA fingerprinting to population biology. This has been brought dramatically to the fore by the recent exploitation by several investigators, using the polymerase chain reaction, of naturally occurring, simple sequence arrays of GT repeats from humans, CAG repeats from *Drosophila melanogaster* and CT repeats from a whale, as hypervariable single-locus probes for easy two-band individual identification in the respective species (see ref. 6 and references therein). Synthetic oligonucleotides based on non-*chi*-

1. Jeffreys, A.J., Wilson, V. & Thein, S.L. *Nature* **314**, 67–73 (1985).
2. Jeffreys, A.J., Royle, N.J., Wilson, V. & Wong, Z. *Nature* **332**, 278–281 (1988).
3. Nakamura, Y. *et al.* *Science* **235**, 1616–1622 (1987).
4. Eickbush, T.H. & Burke, W.D. *J. molec. Biol.* **90**, 357 (1986).
5. Lyons, K.M., Stein, J.H. & Smithies, O. *Genetics* **120**, 267–278 (1988).
6. Tautz, D. *Nucleic Acids Res.* **17**, 6463–6472 (1989).
7. Tautz, D., Trick, M. & Dover, G.A. *Nature* **322**, 652 (1986).
8. Hancock, J.M. & Dover, G.A. *Molec. Biol. Evol.* **5**, 377–391 (1988).
9. Dover, G.A. *J. molec. Evol.* **26**, 47–58 (1987).

Transmembrane protein of SIV

SIR—We would like to add to the observations on the transmembrane protein (TMP) of simian immunodeficiency virus (SIV) reported by Hirsch *et al.* in a previous scientific correspondence¹. Immunoblot analysis of purified virus from a non-cloned pool of infectious SIV_{MAC}, isolated from a rhesus macaque infected with SIV_{MAC251} (virus supplied by R. Desrosiers) demonstrated the presence of a diffuse protein of relative molecular mass 42–46,000 (*M_r* 42–46K) (Fig. 1a). This virus pool, produced by co-cultivation of monkey peripheral-blood lymphocytes with phytohaemagglutinin-stimulated human cord-blood lymphocytes for 17 days followed by growth on the human C8166 T-cell line for 20 days, induces fatal AIDS-like disease in rhesus macaques² and is intended for use as a vaccine challenge stock. Further passage of this virus in C8166 cells resulted first in a population displaying proteins with *M_r* of both 42–46K and 31–37K (Fig. 1c) and later a predominance of 31–37K material (Fig. 1d). Nucleotide sequencing of polymerase chain reaction (PCR) amplified proviral DNA corresponding to the TMP-coding region of the SIV *env* gene revealed the presence of an in-frame stop codon in all six clones derived from the more highly

passed material, whereas none of the eight clones derived from low passage material contained a stop codon within this region (Fig. 2).

We then analysed virus-derived antigen

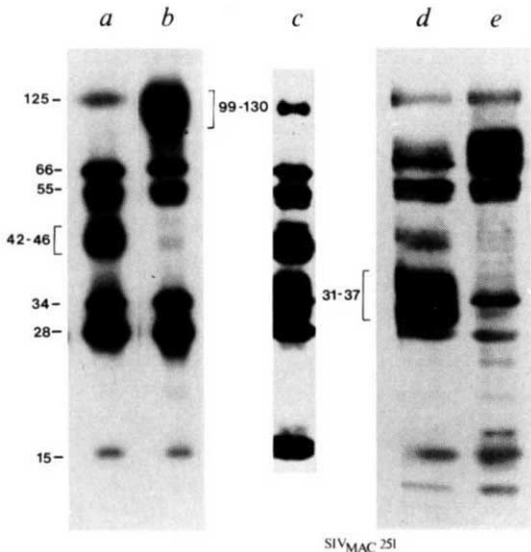


FIG. 1 Virus was purified by ultracentrifugation and chromatography on Sepharose 4b from cultures of infected C8166 cells. The cultures were passed every 3 or 4 days by feeding with uninfected cells in a ratio of 1:3. Tracks a, b, pass 6; track c, pass 14; tracks d, e, pass 18. Tracks b, e were applied in sample buffer without SDS. Virus antigen was electrophoresed on 12.5% polyacrylamide gels containing SDS and proteins detected by probing with serum from an SIV-infected monkey followed by ¹²⁵I-labelled Protein A. Relative molecular masses (K) were estimated from standards.

FIG. 2 Comparison of the DNA and amino-acid sequence of a portion of transmembrane region of the *env* gene of SIV_{MAC}. DNA was isolated from infected C8166 cells at low or high passage. PCR-amplified TMP-coding region was cloned and sequenced by standard procedures and the results compared to those published for two molecular clones of SIV (ref. 6). ▲, Nucleotide change; ■, amino-acid change, ★★★, termination codon.

215	Pro	Ser	Tyr	Phe	***	Gln	Thr	His	Thr	Gln	Gln	Asp	Pro	Ala	Leu	Pro	Thr	231
CCC	TCT	TAT	TTC	TAG		CAG	ACT	CAT	ACC	CAA	CAG	GAC	CCG	GCA	CTG	CCA	ACC	
8771																		8819
32H																		
(low pass)																		
Pro	Ser	Tyr	Phe	Gln	Gln	Thr	His	Ile	Gln	Gln	Asp	Pro	Ala	Leu	Pro	Thr		
CCC	TCT	TAT	TTC	▲	CAG	ACC	▲	ATC	CAA	CAG	GAC	CCG	GCA	CTG	CCA	ACC		
32H																		
(high pass)																		
Pro	Ser	Tyr	Phe	Gln	Gln	Thr	His	Ile	***	Gln	Asp	Pro	Ala	Leu	Pro	Thr		
CCC	TCT	TAT	TTC	CAG	CAG	ACC	CAT	ATC	TAA	CAG	GAC	CCG	GCA	CTG	CCA	ACC		
SIVMAC ²³⁹									▲									
Pro	Ser	Tyr	Phe	Gln	Gln	Thr	His	Ile	Gln	Gln	Asp	Pro	Ala	Leu	Pro	Thr		
CCC	TCT	TAT	TTC	CAG	CAG	ACC	CAT	ATC	CAA	CAG	GAC	CCG	GCA	CTG	CCA	ACC		

blot analysis of purified virus from a non-cloned pool of infectious SIV_{MAC}, isolated from a rhesus macaque infected with SIV_{MAC251} (virus supplied by R. Desrosiers) demonstrated the presence of a diffuse protein of relative molecular mass 42–46,000 (*M_r* 42–46K) (Fig. 1a). This virus pool, produced by co-cultivation of monkey peripheral-blood lymphocytes with phytohaemagglutinin-stimulated human cord-blood lymphocytes for 17 days followed by growth on the human C8166 T-cell line for 20 days, induces fatal AIDS-like disease in rhesus macaques² and is intended for use as a vaccine challenge stock. Further passage of this virus in C8166 cells resulted first in a population displaying proteins with *M_r* of both 42–46K and 31–37K (Fig. 1c) and later a predominance of 31–37K material (Fig. 1d). Nucleotide sequencing of polymerase chain reaction (PCR) amplified proviral DNA corresponding to the TMP-coding region of the SIV *env* gene revealed the presence of an in-frame stop codon in all six clones derived from the more highly

for the presence of oligomeric TMP. It has been reported for HIV-1, that heat stable oligomers of TMP may be detected by immunoblotting when SDS concentrations are reduced in the PAGE sample buffer³. Both full-length (Fig. 1a,b) and truncated (Fig. 1d,e) SIV TMP were found to form such oligomers; the estimated relative molecular masses indicate, in both cases, that these complexes are composed of dimers or trimers or both.

An experimental whole-inactivated-virus vaccine, currently under trial in rhesus and cynomolgus macaques contains full-length and truncated TMP (Fig. 1c). Because infected animals exhibit antibody responses to peptides from the cytoplasmic C-terminal domain of SIV

TMP (ref. 4) and this region modulates growth properties of the virus, at least *in vitro*⁵, we believe inclusion of this region in putative vaccines to be important.

M. P. CRANAGE

Division of Pathology, PHLS Centre for Applied Microbiology & Research, Porton Down, Salisbury SP4 0JG, UK

N. ALMOND

A. JENKINS

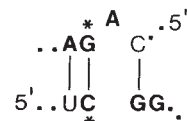
P. A. KITCHIN

WHO AIDS Collaborating Centre, National Institute for Biological Standards and Control, South Mimms, Potters Bar, Herts EN6 3QG, UK

Genetic code origins

SIR—We wish to point out a remarkable circumstance associated with the recent characterization of the guanosine/arginine-binding site in *Tetrahymena* intronic RNA.

F. Michel and colleagues¹, elsewhere in this issue, show that nucleotide pair G264·C311 contributes to the specificity of this part of the catalytic site, and suggest that G264 is hydrogen-bonded to the guanosine/arginine ligands. The remarkable circumstance is that this sidechain- and stereo-selective arginine site² is formed by the conjunction of two arginine-encoding triplets, AGA and CGG:



The base pairs marked with vertical bars form the end of the conserved helix, conventionally called P7, and the arginine codons are in bold type. The critical base pair is between the asterisks.

Arginine has six possible codons, and arginine codons comprise 6.1 per cent of the contiguous triplets in the *Tetrahymena* intron. Given an arbitrary choice of two nucleotides, the probability that both lie within a codon for arginine (in any phase) is 0.029. The above observation would not, therefore, be sufficiently rare in itself to imply more than coincidence. In fact, the CGG codon itself is easily dismissed as coincidental. It is not phylogenetically conserved, instead being rather rare among the 66 tabulated group I intron sequences³.

The AGA codon, however, is different. Although the P7 region is well conserved, there is considerable variation at the AGA codon. Among the 66 group I sequences there are 32 AGA codons (48 per cent), 31 CGA codons (47 per cent), and 3 AGG codons (5 per cent) occurring at the same positions in the sequence. However, these three sets of triplets all encode arginine. Thus, although there is variation on both sides of the completely

- Hirsch, V. *et al.* *Nature* **341**, 573–574 (1989).
- Cranage, M.P. *et al.* in *Proceedings of the International TNO Meeting on Animal Models in AIDS* (ed. M. Horzinek) (Elsevier/North Holland Biomedical, in the press).
- Pinter, A. *et al.* *J. Virol.* **63**, 2674–2679 (1989).
- Franchini, G. *et al.* *AIDS Res. hum. Retrovir.* **4**, 251–258 (1988).
- Chakrabarti, L., Ermerman, M., Tiollais, P. & Sonigo, P. *J. Virol.* **63**, 4395–4403 (1989).
- Naidu, Y.M. *et al.* *J. Virol.* **62**, 4691–4696 (1988).

conserved G264 residue, this variation produces triplets that represent both sets of modern arginine codons, AG $\hat{a}$ and CGN.

All self-splicing group I RNA's so far examined have an arginine-binding site at this same location⁴. Sequence variations flanking the G residue that is apparently in contact with site-bound arginine produce three arginine codons and no others. The unbroken prevalence of this pattern in modern RNAs indicates that the sequences at this point must have also been arginine codons in ancient group I RNAs. This supports the suggestion⁴ that this specific RNA-amino acid interaction may have

been a progenitor of the genetic code. The code, on this basis, probably developed from an ancient RNA codon-amino-acid interaction, as somewhat originally suggested by Woese⁵.

MICHAEL YARUS
ERIC L. CHRISTIAN

*Department of Molecular, Cellular
and Developmental Biology,
University of Colorado at Boulder,
Boulder, Colorado 80309-0347, USA*

1. Michel, F. *et al.* *Nature* **342**, 391–396 (1989).
2. Yarus, M. *Science* **240**, 1751–1758 (1988).
3. Cech, T.R. *Gene* **73**, 259–271 (1988).
4. Hicke, B.J. *et al.* *EMBO J* **8**, 3843–3851 (1989).
5. Woese, C.R. *et al.* *CSH Symp. quant. Biol.* **31**, 723 (1966).

Homoeopathic test

SIR—The practice of homoeopathy depends on the use of medicaments prepared in a special way. Solutions derived from a variety of parent substances are progressively diluted, with a dilution factor of 1 in 100 at each stage. The resulting fluid is said to acquire a specific 'potency'. In clinical use it is essential that the medicament chosen shall be appropriate to the patient's condition; fundamental to this is the contention that the various potencies can be distinguished from one another. Some years ago, I put that contention to the following test.

In consultation with an eminent homoeopathic practitioner, a former

ment should be placed in each bottle. Each bottle was enclosed in a metal can and the whole batch was sent to the practitioner without revealing the key.

No constraints were placed on the practitioner's methods, whether chemical, physical, clinical, parapsychological, or even magical, for identification of the potencies, as the question at issue was whether any distinctions could be detected regardless of means. But the practitioner kept to the routine he used in clinical practice for matching remedies to ailments. In all, he carried out 19 complete and reputedly independent tests on the same set of 20 bottles between April 1958 and May 1967, and reported his final conclusions in March 1970.

TABLE 1 Hypothetical and measured distributions from homoeopathic tests

	I <i>n</i> = 19	II <i>n</i> = 20	III <i>n</i> = 380
Calculated			
1. <i>p</i> = 1.0	19.0	20.0	380
2. <i>p</i> = 0.9	17.13 ± 1.3	18.0 ± 1.34	342 ± 5.85
3. <i>p</i> = 0.7	13.3 ± 2.0	14.0 ± 2.05	266 ± 8.93
4. <i>p</i> = 0.6	11.4 ± 2.13	12.0 ± 2.19	228 ± 9.54
5. <i>p</i> = 0.5	9.5 ± 2.18	10.0 ± 2.24	190 ± 9.22
Measured	10.63 ± 2.14	10.1 ± 2.55	202

TABLE 2 Results of comparisons using Student's *t*-test

	I <i>n</i> = 19	II <i>n</i> = 20	III <i>n</i> = 380
1. <i>p</i> = 1.0	reject	reject	reject
2. <i>p</i> = 0.9	reject	reject	reject
3. <i>p</i> = 0.7	<i>t</i> = 3.9 (reject)	reject	reject
4. <i>p</i> = 0.6	<i>t</i> = 1.09	<i>t</i> = 2.49 (reject)	<i>t</i> = 2.72 (reject)
5. <i>p</i> = 0.5	<i>t</i> = 1.59	<i>t</i> = 0.13	<i>t</i> = 1.3

President of the Faculty of Homoeopathy, two specific potencies were selected for the test, namely, 'Natrium muriaticum 30C' and 'Sulphur 30C'. They are said to have strikingly different properties and to be strongly active. The practitioner declared himself quite confident that there could be no possibility of confusion between them.

A set of 20 uncontaminated bottles were numbered and then loaded with one or other of the medicaments, using a set of random numbers to decide which medica-

The results were analysed in three ways: I, the number of correct identifications per set of 20 different bottles examined in a single run; II, the number of times each bottle was correctly identified in the 19 attempts on the bottle; III, the total number of correct identifications overall.

These scores are compared with those expected on the following hypotheses: 1, that the potencies are identifiable with certainty (that is, probability *p* = 1) as originally envisaged by the practitioner; 2, 3, and 4, that a correct identification could

be expected in 90, 70, or 60 per cent of trials respectively (*p* = 0.9, 0.7 or 0.6, respectively); 5, that the observed number of correct identifications could have been arrived at by some random procedure which carries no information about the nature of the items (*p* = 0.5).

Table 1 sets out the means and standard deviations expected under each of these 5 hypotheses, for each of the 3 groupings of data, together with the actual values found. Table 2 sets out the conclusions from the various comparisons on the basis of Student's *t*-test at the 5% level.

The result for I (*n* = 19) cannot be distinguished either from the distribution for *p* = 0.6 or from that for *p* = 0.5 (chance). The other comparisons lead to the rejection of all the hypotheses that involve some discriminative power, leaving only hypothesis 5, indicating that the results might just as well have been arrived at by chance alone.

For a truly blind trial, each bottle should have contained material with a variety of possible activities or no activity at all. The practitioner may have been influenced by the expectation that each bottle should show activity of one or other type, in equal numbers and with no blanks, thus making the task of identification easier. Despite this bias, the statistical tests do not support the claim that the two potencies can be distinguished, at least by the methods used by this particular practitioner. It is possible that such a distinction could be derived from their clinical use, but practitioners are understandably reluctant to carry out a double-blind clinical trial that involves giving a human patient the wrong medicament. Perhaps a veterinary practitioner might be less reluctant.

T. D. M. ROBERTS

11 Menzies Drive, Fintry,
Glasgow G63 0YG, UK

Amazon flooding

SIR—In his News and Views article on Amazonia during the ice-ages, P.A. Colinvaux referred to my hypothesis that the Amazon Basin was flooded to the 250–300-m contour by a great freshwater lake during the last glacial from about 45,000 BP (but not during previous glacials as might have been understood from the text). This lake, of course, was held in by a dam across the Amazon River, not by dams across its tributaries, which were also flooded, as Colinvaux stated. As well as draining north through the Orinoco, the lake would have drained south through the llanos of Bolivia to eventually reach into the South Atlantic.

KENNETH E. CAMPBELL

Natural History Museum of
Los Angeles County,
900 Exposition Boulevard,
Los Angeles CA 90007, USA

A turn in the weather

Peter J. Parr

Out of the Desert? Archaeology and the Exodus/Conquest Narratives. By William H. Stiebing, Jr. *Prometheus*: 1989. Pp. 265. \$21.95, £16.50.

OLD Testament history used to be relatively simple. The Bible was, after all, generally recognized as the most complete and detailed — as well as the most readable — historical source to have come down to us from the pre-classical world, and one did not have to be a fundamentalist to believe, more or less, what it said.

But then, as a result of the literary criticism of the early part of this century, it became clear that the historical books of the Old Testament were far from being a unified narrative; rather, they were a mosaic of documents that had been written at various times, often centuries after the events they purported to describe, and had been edited with varying degrees of scholarship by a number of different people on a number of different occasions. The burgeoning archaeological discoveries in Palestine during the middle decades of the century

did little to improve matters. After some initial optimism that the spade was beginning to uncover undoubted traces of the events of biblical history, it emerged that archaeology was as unsatisfactory and tantalizing a guide as the written word. The result is that for almost every aspect of the history of ancient Israel, there is now a wide range of often totally contradictory hypotheses available. All of them are capable of explaining at least some of the supposed facts; but none is convincing enough to command universal, or even majority, respect.

Of nothing is this more true than of the events that lie at the heart of Israel's history, and are the foundation of its religion and its national identity: its occupation of the land of Canaan and its dispossession of the previous inhabitants (the Exodus and Conquest). A vast literature has grown around the problems of chronology and causation surrounding these events. Most scholars have their own individual theories, but, as Professor Stiebing shows in the excellent volume under review, there are four or five basic hypotheses.

The first, and longest established, is the most conservative (with the exception of the real fundamentalist approach), and accepts the fact of an Israelite exodus from Egypt followed by an essentially forceful subjugation of Canaan, though not necessarily in all the particulars given in the biblical account. The second is of

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

On the go — the children of Israel crossing the River Jordan.

almost equally long ancestry. It also bases itself fairly closely on the biblical text, but sees the settlement of the Israelites more as the result of peaceful infiltration into the less densely inhabited areas of Canaan by semi-nomadic pastoralist groups from the desert fringes. A third model, proposed in the early 1960s, has proved popular with historians adopting the fashionable socioeconomic stance. It postulates an entirely internal development within Canaan, a sort of 'peasants revolt' against oppression by the ruling classes and the flight of disaffected groups away from the urban centres and towards more remote parts of the land, where their relative isolation and sense of common grievance and purpose fostered their eventual growth into a nation.

Finally, there are what Stiebing calls composite or hybrid theories which have developed from the other models. One of these sees the original Israelites not as social and political revolutionaries, but as refugees escaping from the onslaught of more warlike neighbours, such as the Philistines; another accepts the semi-nomadic pastoralist origin of the Israel-

ites, but argues that they were not newcomers from outside the country but had been in Canaan for generations, living in symbiosis with the town-dwellers and becoming sedentary only when the Canaanite urban centres on which they depended collapsed. Just when any of these processes or events took place is also a matter of controversy; current suggestions range between the twenty-third century BC to the eleventh, though most serious scholars accept the thirteenth as the most likely, whichever theory is adopted.

By far the greater part of this book is devoted to a discussion of the arguments for the various theories and dates, each of which has its strengths and weaknesses.

The task has been done many times before, but rarely as lucidly, succinctly and judiciously as by Professor Stiebing. After an admirably straightforward account of the principles of biblical literary criticism and a review of the relevant Palestinian archaeological material, the evidence for and against each hypothesis and each suggested chronology is presented fully and fairly, and is critically assessed. None of the theories is found to be wholly satisfactory, and it is this somewhat negative and disappointing state of affairs that leads Professor

Stiebing to an eminently sensible conclusion, namely that "one problem with all of the models we have considered is that they view the Israelite conquest or settlement in relative isolation". As he points out, the final centuries of the second millennium in the eastern Mediterranean and western Asia saw some of the most dramatic and turbulent events of ancient history: the collapse of the Hittite, Egyptian and Mycenaean empires, and the appearance of the Dorian Greeks, the Sea Peoples and the Aramaeans, as well as of the Israelites.

Were all these events connected; did they have a common cause? Stiebing thinks they were, and argues that it was widespread drought resulting from climatic change which led to the economic, social and political disturbances throughout the region. The argument is not new, and the entire subject of the influence of climatic change on the course of human history is one which has for long engendered heated debate between determinist and anti-determinist schools of historical interpretation. The evidence from the climatological and geological sources which Stiebing reviews — though certain-

The Mansell Collection

ly less fully and critically than he does that from archaeology and the texts — is not as conclusive as he implies, if only because there is still far too little of it. But the argument remains attractive and intellectually satisfying, even if intuitive, and helps establish an 'ultimate cause' of the Exodus and Conquest.

The final question, however — which Stiebing does not address — is whether such an 'ultimate cause' really matters. A distinguished historian, Geoffrey Elton, once said that his colleagues tended to spend too much of their time on causes and not enough on results. Climatic change is a phenomenon of concern to the scientist, but what should interest the historian, and what he should seek to

explain, are the different reactions of different people to the same challenge. The disturbances at the end of the second millennium may well have their origin in the same climatic events, but the consequences in various parts of the Mediterranean, Western Asia and Egypt were very different. In Canaan they resulted in the creation of the Israelite nation; and the fact that the Israelites believed their nation to have been created by Exodus and Conquest is of much greater interest and importance than whether it was really true. □

Peter J. Parr is in the Department of Western Asiatic Archaeology, Institute of Archaeology, 31-34 Gordon Square, London WC1H 0PY, UK.

More heat for the debate

R. S. Pease

Nuclear Fusion. By Keishiro Niu. Cambridge University Press: 1989. Pp.240. £35, \$59.50.

JAPAN'S research into controlled nuclear fusion became established as a substantial national programme in the 1970s, and involves a remarkable partnership between the universities, the Japanese Atomic Energy Research Institute and engineering industries. About a dozen universities have research programmes; there is a central institute devoted to plasma physics and magnetic confinement at Nagoya, and a leading institute for lasers and inertial confinement at Osaka.

Professor Niu is head of the theory group of the department of energy studies at the Tokyo Institute of Technology, and has contributed to research in fusion reactor studies. His book, which is a graduate-level text for engineering science, was originally published in Japanese in 1979 with M. Sugira of the Electrotechnology Laboratory, Tsukuba, as co-author. Now it has been translated and partially updated by Professor Niu, who presents the principles of controlled thermonuclear fusion in high-temperature plasmas confined by magnetic fields and by inertia.

The merit of the book is the wide range of topics covered. The introductory chapter leads to the Lawson criterion for the degree of confinement required for a net energy output from thermonuclear reactions, given in a form suitable for quasi-continuous systems. A chapter on magnetic confinement covers the basic properties of charged-particle motion in magnetic fields, the concept of magnetic pressure and the various heating methods. Most space is given to the Tokamak — currently the leading magnetic confine-

ment system — and to instabilities which restrict the configurations that can be used, and those which may restrict the thermal insulation. Other magnetic systems are dealt with briefly.

The chapter on inertial confinement describes the basis of the compression and heating of small spheres of deuterium-tritium fuel by focusing onto them powerful beams of pulsed laser radiation and of charged particles. Here the author elaborates on the problems of the absorption of the radiation, describes the lasers and the methods of generating and focusing particle beams. The final chapter, on reactor design, describes some of the engineering issues, touches on resources of materials and safety, and describes three of the reactor concepts that have been developed: for the Tokamak, for mirror confinement systems and for inertial confinement, respectively.

The book deals hardly at all with experimental results, and it is disappointing to have no account of the impressive Japanese achievements in that area. Professor Niu has given himself insufficient space to expound some of the topics adequately — reading the book, one feels rather like a member of a touring party led by a ruthless guide, determined to show us everything, however imperfect our appreciation. He has also been poorly served by the proof-readers. The equation of motion describing the balance of electromagnetic forces with plasma pressure gradients is corrupted by misprints, for example; the reverse-field pinch configuration is said to arise from a diamagnetic effect, whereas it is paramagnetic; and *aficianados* of mistranslation will appreciate the "Miller" field for magnetic confinement described in Chapter 2, but translated correctly as the "mirror" only in the last chapter. Taken altogether, the book is a useful but not wholly accurate guide to the principles of hot fusion. □

R. S. Pease was formerly Director of Fusion Research at the UK Atomic Energy Authority.

Peas before swine

J. S. Jones

The Mendelian Revolution: The Emergence of Hereditary Concepts in Modern Science and Society. By Peter J. Bowler. Athlone, London: 1989. Pp. 207. £35, \$29.95.

THE sociology of science bears the same relation to research as pornography does to sex: it is cheaper, easier and — as it is constrained only by the imagination — can be a lot more fun. This book discusses the changes in our thinking about heredity and evolution over the past 200 years. Those of us who teach genetics like to present the subject as unfolding logically from Mendel, through Morgan, to Crick and the latest heart-throb of molecular

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Mendel's work was thought 'irrelevant' biology. We all accept that there may be some self-delusion here, but it makes us feel good about the rigour of our own field, and it keeps at least a few of the students awake.

Bowler sees this as a conspiracy by geneticists to hide the subject's past. In this slim volume (and at a price which will make its purchasers expect something Rather Special) he argues that scientific progress makes sense only in its historical context, and that this is particularly true of the study of inheritance because of its use in justifying social policy. As he says, "most historians have little interest in defending the old view that science offers

completely objective knowledge". In pursuit of this thesis, Bowler produces some interesting insights into early genetics. Few of us now realize just how unimportant the mechanism of inheritance appeared to nineteenth-century biologists. They were much more interested in cells, in development and in the nature of species, and what we now call genetics was seen as a minor part of each of these fields. Mendel's work was not ignored because it was not understood; rather it was thought to be irrelevant. Modern genetics could not begin until the study of the transmission of biological information was separated from that of its translation, and this book contains a great deal that will change the views of geneticists about the beginnings of their subject.

It is less satisfactory on the evolution of genetics in the twentieth century. Many would quarrel with the claim that Morgan became interested in the chromosome theory because it gave him access to funds, and that in the 1950s to study cytoplasmic inheritance was to side with communism. And is it really true that Kammerer's work on the inheritance of acquired characters was never duplicated only because he had an unrivalled ability to breed midwife toads in captivity? Bowler ends by pointing out the irony that molecular biology has now turned the study of inheritance full circle by casting doubt on the concept of the gene as a unit.

The book sees genetics primarily as the raw material of history and philosophy. It is a study of science and society and, of its kind, a good one. No doubt in a hundred years we will have similar treatments of selfish DNA as an example of Thatcherite greed, or of genetic imprinting and the women's movement; and there will arise new Poppers, Kants and Kuhns — a sort of philosophical Joy of Sects — all anxious to give biologists the instruction manual of *How Science is Actually Done*. On the tomb of a philosopher buried just across the path from Herbert Spencer (who makes an unexpected appearance in these pages as a social Lamarckist) is the epitaph "The philosophers have only interpreted the world in various ways. The point, however, is to change it". Alter the object of the first sentence to 'science' and most geneticists will agree with Marx's interpretation of their intellectual priorities. □

J.S. Jones is in the Department of Genetics and Biometry, University College London, 4 Stephenson Way, London NW1 2HE, UK.

• *The Extended Phenotype*, by Richard Dawkins, has recently been reissued in paperback by Oxford University Press. The book is the sequel to *The Selfish Gene*, and was reviewed in *Nature* 296, 506 (1982). Price £6.95.

• *Toward a New Philosophy of Biology* by Ernst Mayr is also now out in paperback. Published by Harvard University Press \$14.95, £11.95. See review in *Nature* 333, 25 (1988).

Drawing back the cloak

Ralph A. Lewin

The *Chlamydomonas* Sourcebook: A Comprehensive Guide to Biology and Laboratory Use. By Elizabeth H. Harris. Academic: 1989. Pp. 780. \$145, £90.

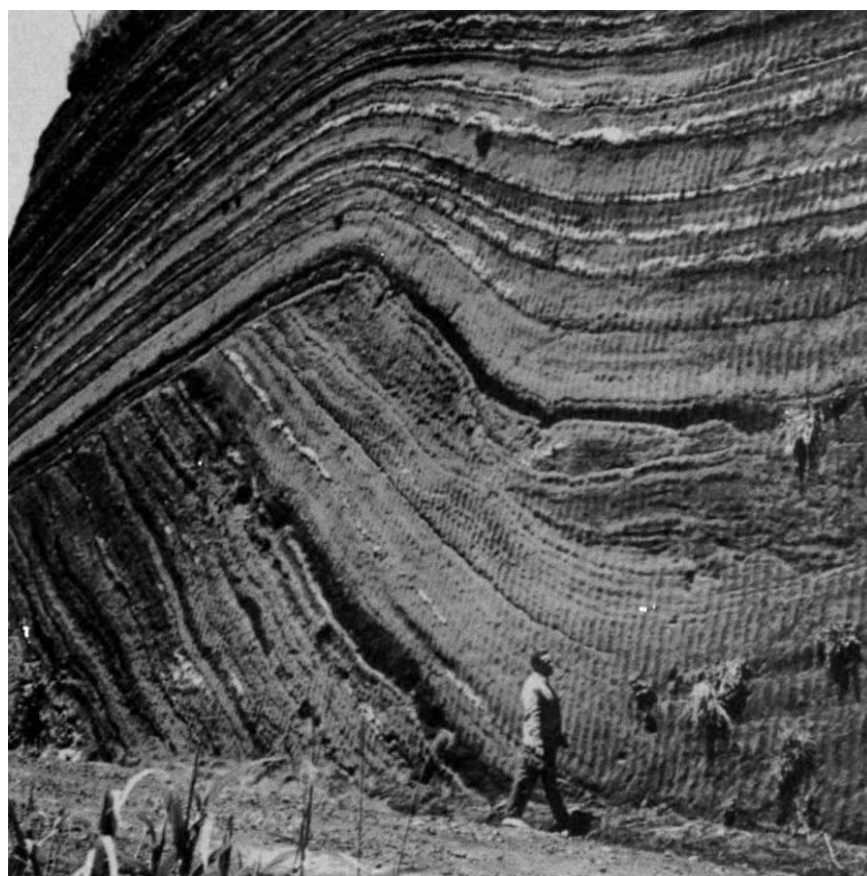
THE name *Chlamydomonas* — from the Greek, 'the cloaked one' — suggests something melodramatic, mysterious, perhaps even furtive. But when Ehrenberg introduced the name in 1833 he could have had no idea how much melodrama would emerge from the study of this simple-looking, single-celled green alga, or how much mystery would remain even today.

The genus is a large one, comprising many hundreds of species: a few marine, more in freshwater ponds and streams, and lots more in soils, even in deserts, from polar regions to the tropics. Put a teaspoonful of soil into a glass of water, add a pinch of fertilizer, set the glass in a window, and in a couple of weeks you'll assuredly have a culture with one or more kinds of *Chlamydomonas*, their micro-

scopic cells swimming by means of paired anterior flagella.

To identify them further a reference work will be needed. So far, the best available has been a detailed taxonomic compilation in German, by H. Ettl (*Nova Hedwigia Beih.* 49, 1–1122; 1976). Now there is an equally scholarly and valuable book to tell us almost everything else that we need to know about the genus — or, rather, to outline what has been discovered, how and by whom, what still needs to be studied, and how we should set about it. The scope of *The Chlamydomonas Sourcebook* is vast, but Libby Harris has been equal to the task. We have here a paragon of excellence.

The book presents, in clear, readable English, an overview of the huge and (almost certainly) polyphyletic genus, with chapters on cell structure, motility (those two simple-looking flagella have proven to be fantastically complicated machines), photosynthesis and other metabolic processes, and genetics of nuclear and non-nuclear genes (chiefly those of the chloroplast, now probably the best known in all biology). There are also detailed instructions for the cultivation, mutagenesis, mutant selection and conservation of strains along with some very useful



Volcanic ash layers on a road cut in Oshima Volcano, Japan. Airfall material, from volcanic eruption, forms distinct layers that follow the slope of the ground. Two generations of ash fall are seen in the picture, the beds are not folded; the inclined layers reflect the original attitude of ash accumulation. The photograph is reproduced from *Volcanoes*, by R. Decker and B. Decker, a revised and updated edition published by W. H. Freeman. Price is £11.95, \$14.95.

information on the ones in current use. (As in the Bible, there is a 'begat' chapter with all sorts of details of family histories, carefully researched by Harris before memories fade and sources are forever forgotten.) Also included are descriptions of many useful biochemical techniques and methodologies for molecular biological work on such unicellular algae.

There are, too, good reviews of the sexuality of these versatile microbes, the feature which, above all else, led to their selection for detailed studies. Seventy years ago, the first direct confirmation of mendelian inheritance came from a study of *Chlamydomonas*, but for 30 years after that there were few publications on this subject (and many of those, regrettably, proved to be largely fictional). Only in the late 1940s did a handful of biologists, seeking a single-celled plantlet with genetic potential, take up studies of the genus. Mostly we chose to work with one of two unrelated species, *C. moewusii* (also known as *C. eugametos*) and *C. reinhardtii*. Their sexual proclivities are sometimes melodramatic, or even furtive, but by now they have been elucidated fairly well and are under good laboratory control. Unfortunately *C. moewusii*, which lends itself best to studies of sexuality, suffers from a metabolic disability all too common among algae, an inability to grow heterotrophically in darkness, so in this species photosynthesis cannot be bypassed. Consequently *C. reinhardtii*, with its more versatile metabolic potentials, has been the species of choice.

If I have to find a bone to pick, it would be on this point. When you read in Harris's book something about *Chlamydomonas*, you can generally assume that it relates to *C. reinhardtii* unless otherwise specified. However, generalizations are dangerous in a 'genus' as diverse as this — what's true for Peter may not be applicable to Paul — and doubtless examination of other species will reveal exceptions and surprises of all sorts.

I note that on p.175 the first equation, misquoted from Chiang *et al.*, is incorrect, and that on p.130 *C. philotes* is wrongly given as heterothallic. But I have discovered nothing of importance that's missing, and lots of fascinating details that I knew nothing about before. There are full references to almost 3,000 publications, and abundant tables, diagrams, photographs and other illustrations, with adequate scales. A good ten-page, three-column index is provided. It's all there, in a stout binding and a hard, waterproof cover. The book will inevitably need both. □

Ralph A. Lewin is in the Marine Biology Research Division, Scripps Institution of Oceanography, University of California San Diego, La Jolla, California 92093, USA.

Far off and exceeding deep

Harry Shipman and Noel Hart

Deep Time: The Journey of a Single Subatomic Particle from the Moment of Creation to the Death of the Universe and Beyond. By David Darling. Bantam, London/Delacorte, New York: 1989. Pp.192. £12.95, \$17.95.

From Quarks to the Cosmos: Tools of Discovery. By Leon M. Lederman and David N. Schramm. W.H. Freeman: 1989. Pp.242. £16.95, \$32.95.

How did the Universe begin? Where will it end? These questions have taken on a new, rich dimension in the 1980s with the development of a connection between cosmology and particle physics. In producing these two books, Darling and Lederman and Schramm join the crowd of scientists and journalists writing on the subject for a 'popular' audience (whatever that is).

Deep Time is a 'prose poem' telling the story of a proton which journeys from the Big Bang to the end of cosmic evolution. Some episodes in this journey illustrate a different and entertaining way of describing astronomical phenomena. Our proton hero narrowly escapes being sucked into a black hole in the centre of a galaxy, and along the way there is a nice description of the current central dogma about quasar energy sources.

But any writer takes on a real challenge in writing a book-length 'poem'. Lewis Carroll pulled it off in *Alice in Wonderland*, but Darling overdoses the reader with phrases such as "the hot breath of genesis", "oases of order" and "the mighty fortress of the second law" (of thermodynamics). One of us (H. S.) found the language merely bothersome; the other (N. H.) found it virtually unreadable.

It is hard to draw distinctions between well-founded theory and scientific speculation in a poem. Darling doesn't bother. For instance, he implies that cosmic strings must be midwives that assist in the birth of galaxies. Yet we don't even know whether cosmic strings exist. Fiction comes in, too, with tales of Prospero, the tenth planet in the Solar System. What tenth planet? Inexpert readers are here being led into the land of the tabloids, where the difference between science, pseudoscience and science fiction is considered to be irrelevant. Such readers will also have to struggle with sentences that presume a more than passing familiarity with, for instance, antileptons and antibaryons.

Lederman and Schramm provide a more conventional treatment of cosmology and its interface with particle physics.

The pictures are outstanding and serve well to illustrate the sparse text, while Lederman's perspective as an experimental particle physicist shows through nicely. Many pages clearly and thoroughly, though ponderously, describe mammoth particle accelerators and the associated array of detectors. Most popular books on this topic let the stork deliver new particles, such as the omega-minuses, from the clear blue sky. The authors here give us the facts of life, describing the 50,000 pictures which had to be taken before the omega-minus appeared to confirm the symmetric view of nature based on the quark model.

Lederman and Schramm, however, use jargon injudiciously. The book's second sentence describes the control room of Fermilab's Tevatron, but those who don't know what or where the Tevatron is must wait until page 115 for a pleasant, four-page description of it. And the beat goes on: in the first five pages the reader is sprayed with bullets of terminology: anti-protons, p-bars, accumulator storage rings, luminosity, deflection 'kicker' magnets, GeV, CDF, standard model, quarks, leptons and more. One-sentence definitions of some of these bits of scientific lingo are not enough for the non-specialist. Several pages later the level plunges precipitously. Using a box, the authors patiently explain powers of ten, and they take ten pages of text and many gorgeous illustrations to show that atoms are made of electrons and protons. Both reviewers, one of us a scientist and one a non-scientist, found the text to be like a rollercoaster ride and just as disorientating.

There is a plethora of cosmology books on the market already, including, of course, Hawking's blockbuster *A Brief History of Time*. Why write or publish so many of them? That's a question the book-buying public should be asking. □

Harry Shipman is in the Department of Physics and Astronomy, and Noel Hart the Department of Foreign Languages and Literatures, University of Delaware, Newark, Delaware 19716, USA.

● An English translation of Paolo Maffei's *L'Universo nel tempo* is now available; *The Universe in Time*, published by MIT Press, addresses issues not covered in Maffei's first book, *Beyond the Moon*. *The Universe in Time* is structured as a voyage in time, looking at the life cycle of the stars, planetary systems and eventually geological and ecological evolution. Price \$29.95, £26.95.

● Also recently translated, from the French, is Jean Heidmann's *Cosmic Odyssey*. The book was written to "give a description of the fabulous richness of the cosmos". It explores the relativistic universe, the quantum universe and the inflationary universe. *Cosmic Odyssey* is published by Cambridge University Press, price £11.95, \$19.95.

Brazil walks the tightrope

Brazilian researchers survive with admirable calm amid an economic tempest. With strong nerves and good luck, the first democratically elected president in decades may next year bring better times.

THE first round of elections to choose Brazil's president has presented the country with its first true choice since the military takeover in 1964. In last week's free and democratic vote, two candidates with diametrically opposite views were chosen to go forward to the final vote on 17 December. Both candidates are young but their backgrounds and outlooks are utterly different.

Fernando Collor de Mello (40) is wealthy, handsome and conservative. His main, perhaps only, claim to fame is the action he took to cut the salaries of 'maharajas' — enormously overpaid government officials — while governor of the small north-eastern state of Alagoas. His National Renovation Party was set up simply to promote him and has no ideological or class base. Facing him is Luíz Inácio da Silva, known simply as 'Lula', a former lathe operator and leader of the autoworkers' union. Almost exactly ten years ago he led a strike of three million workers that opened the first cracks in military rule.

His Workers Party is well-organized, well-established and supports a redistribution of income, regulation of capitalist enterprise and a moratorium on the repayment of the foreign debt.

So far there has been no sign that the military will intervene in the electoral process unless forced to do so by a major social upheaval. But if Lula is elected next month, the temptation may be strong, and there will be anxiety in Brazil. One factor staying the military's hand will be the knowledge that they have no idea how to tackle Brazil's colossal economic problems.

The best the winner can hope for when he takes over from President José Sarney in January is that the economic condition of Brazil will be no worse than today. Brazil is close to terminal hyperinflation and economic disintegration. In the five years during which Sarney has been president, nothing has been done to tackle the underlying problems.

Last year, inflation came close to 1,000 per cent. Now it is running twice as fast, at over 30 per cent a month. In daily life, the evidence of inflation is immediate and tangible. Take the 100,000 cruzeiro note; in 1986 it was worth almost ten dollars, the biggest note in common use. Three years later, the note comes as change for the smallest of purchases: it is worth less than one cent.

Money is not something solid in Brazil, it is constantly slipping away and vanishing. Fixed salaries are halving in value in less than two months; unpaid bills a few months old are not worth collecting, and items ordered a few months ago will double or triple in price before they arrive. Brazilians never know how much anything costs — only what it used to cost. Restaurants do not print prices on their menus but just leave spaces to be filled in anew each day. Taxi drivers have their meters set at some historical rate and then consult an official table — which always appears freshly issued — to convert the reading to today's charges.

Inflation has left a bewildering trail of banknotes. The 'cruzado' was created in 1986 in the Cruzado plan, the first of the government's two big attempts to halt inflation. It is worth 1,000 of the old cruzeiros. The next attack on inflation, the Summer plan, created the 'new cruzado' worth 1,000 of the old cruzado. Notes of all three currencies are in circulation, meaning that the face value of some has to be divided by a thousand and others by a million.

Dealing with inflation obsesses everyone. Salaries are constantly being revised and re-revised to try to keep up, Congress is always tied up with supplementary budgets to restore some value to budgets granted earlier. Anyone with any significant sum of money rushes it — immediately — into high-interest accounts. It is

With thanks

One person and one person alone made possible the four-week trip to Brazil upon which this survey is based. Ricardo Bonalume Neto, a science journalist at the newspaper *Folha de São Paulo* and occasional contributor to the news pages of *Nature*, kindly volunteered to set up appointments for me within Brazil, a task which is made triply difficult by frequent breakdowns in Brazil's long-distance telephone system. Despite the problems, Ricardo always succeeded in reaching the unreachable and making appointments with the totally unavailable — and always with just moments to spare. In São Paulo, special thanks go to Ricardo for driving me around town in his vintage Beetle, despite a series of punctures.

Alun Anderson

possible to make almost two per cent a day, 55 per cent a month. Industry can deal in special inflation-indexed Treasury bonds and keep itself afloat. But for those on fixed incomes, and for the poor, it is a different matter.

Brazil contains vast inequalities. In the wealthy area behind Copacabana beach, every apartment building has its guard. Late at night they are the only people on the streets. Not far away one million people live in *favelas*, shanty towns — they too are visible from the beach. In São Paulo, shacks of wood and cardboard are crammed under elevated motorway inter-sections and line the abandoned land along the city's polluted rivers. Altogether 25 million Brazilians live in slums out of a population of 130 million.

The statistics of inequality are staggering. The top five per cent of the population control almost 50 per cent of the nation's wealth, the next 25 per cent take 30 per cent; the remaining 70 per cent of the population share 20 per cent. More than 20 million people do not have an adequate diet, five million children have never attended school, 20 per cent of all adults are completely illiterate and fewer than half can read a newspaper.

But then comes the great contradiction. Despite its third-world distribution of wealth, Brazil does not have a third-world economy. Brazil is the world's eighth largest economic power, with a GNP of \$250,000 million and a \$20,000 million trade surplus — the world's largest after Japan and West Germany — on exports of \$33,000 million. And those exports are not the agricultural products and raw materials of a third-world nation; 70 per cent of them are manufactured products including aeroplanes, automobiles, industrial machinery, tanks, and even missiles.

Thanks to the power of the economy, the top quarter of the population — 30 million or so people — are middle class with much the same tastes in clothes, food and culture as their European or American counterparts. In the book shops in São Paulo, the bestsellers are the same as in New York — Umberto Eco, Milan Kundera and even Stephen Hawking's *Uma breve historia do tempo*. Look one way and Brazil is part of the developed world — look the other way and it is part of the third world. Brazil is in transition. What of the future? Is Brazil about to make the leap to developed-nation status or fall back to third-world status?

If Brazil is to emerge from its difficulties it will have to tackle its huge deficit. It is doubly hampered; it owes a huge amount of money — over \$100,000 million — to foreign bankers and it has a colossal internal debt. The oil money borrowed during the 1970s was poured into vast state-owned enterprises: oil, steel, chemicals, electricity, railways and telecommunications. State capitalism may have kick-started Brazil's development but it never proved much good at providing profits. Today, the state enterprises are a huge drain on government resources. They have never been left to follow market realities but are used for a variety of political purposes. When convenient, electricity prices are set to manipulate the rate of inflation or jobs are given out to relieve unemployment in a sensitive area.

The need to finance the debt of the state enterprises is one of the factors that triggered inflation. As foreign loans vanished, the government turned to the internal market or printed money. And as the government is the guarantor of huge quantities of money in index-linked savings and bonds, it had to borrow to pay for inflation, driving the inflation rate yet higher.

The state enterprises are a social as well as an economic phenomenon. They created a new breed of Brazilian — the technocrats. Well-paid and enormously powerful, they will prove reluctant to work anywhere else. So will the army of workers they control. For all but the richest, Brazilian life is permeated with a fear of falling — falling into the abject poverty below. In a state sector job security is total — and wages are higher. Every year more jobs are created and labyrinthine bureaucracy grows.

It will be a brave politician who tackles the state enterprises. Sarney made one brief sally in January when he announced plans to dismiss 60,000 state employees (one result would have been that almost half the staff at the research facilities described below would have been out of a job). But he backed down in the face of public opposition, saying the matter was the responsibility of Congress.

What will Collor or Lula do? Collor has so far been careful to cultivate his image while not committing himself to any definite policy, although he has said he is in favour of the sale of state enterprises. Lula's views are clearer. Corruption and nepotism are to go from the state enterprises but the enterprises themselves will remain and perhaps even be strengthened.

Without tackling the deficit it is difficult to see how the Brazilian economy can recover. But there are other things that can be done to help, Bolivia providing one example. In 1985, Bolivia's inflation rate reached 23,000 per cent but now it is under 10 per cent. Hyperinflation was ended by getting rid of the deficit, in part by the

unorthodox method of reducing tax rates to the level where people — particularly those operating in the black economy — would pay them rather than cheat. Tax revenues rose several thousand times as a result. In Brazil, some economists believe that 50 per cent of the economy may be black. If that is the case, then changing the tax structure to bring them into the legitimate economy, and removing excess regulation that makes it hard for them to expand, could provide the boost Brazil needs.

Despite the sense of impending economic catastrophe, the visitor is surprised to find that Brazilians themselves are rarely downcast. Amid all the difficulties, it would be hard to find a Brazilian who does not believe the nation is destined for greatness. They may joke that "Brazil is the country of the next century — and always will be", but underneath there is unfailing optimism. They know that the economy can be dynamic, they know that the country is vast and vastly rich with minerals, oil and gas. When pressed, that confidence slips into the kind of xenophobia which inspires politicians and military men to threaten to cancel the foreign debt and to tell the rest of the world to go to hell.

It is in this spirit that Brazil's politicians talk about the Amazon — its their great frontier, the American West, that will be conquered and bring untold riches. The development of the Amazon is inevitable; the only questions are when and how.

But international pressure has put the environment on the political and election agenda. Some tax subsidies for 'development' of the Amazon that have encouraged giant cattle ranches have already gone, and if the pressure is kept up, others may go too. But politicians know that the main force driving poor settlers to Rondonia is poverty; expansion in the industrial heartlands would be the best hope for the Amazon, buying time until better development strategies have been worked out. But to say that is simply to return to the problem of inflation.

In the pages that follow, it will be readily apparent that scientists have no escape from the economic, political and social realities that surround them. Scientists worry about inflation, financial instability (which sends some institutes close to bankruptcy), sudden reversals in policy and the wave of democracy which is altering old power structures within research institutes. It is a time of transition. Among the cast of characters that appear in the remainder of this survey are the giants of the older generation, people who weathered the period of military rule or who had to flee into exile, a new breed of research administrators who know the routes through the labyrinths of Brasilia, frustrated young researchers fresh from foreign research laboratories, and those

who have put aside personal goals to train the next generation of researchers and work for development. And of course there are some bureaucrats and technocrats too.

People will be found to disagree on many things, as may be expected. The reader must also be warned that not everything told to *Nature* is necessarily true. In Brazil, charm is a national art, and even scientists may sometimes be unwilling to let facts interfere with a well-told tale. Charm is necessary. Amid Brazil's inequalities it makes life bearable, even enjoyable. At one end of the spectrum, charm blends into the kind of good-natured, smiling persistence that is often necessary to get anything done in Brazil. Further along it turns into the skill of creating imaginative compromises and ways around rules (the '*jeito*' of which Brazilians are so proud). Yet further along it turns into corruption plain and simple.

In the opinion of this reporter, one category of researchers is special. Those who live and work in the Amazon are set apart by the extraordinary paucity of their resources and by the unique interest of their research into one of the world's least understood ecosystems. Brazil could buy some very inexpensive international prestige by deciding to make research in the Amazon a high priority — after all, the Amazon (or the greater part of it) is Brazil's alone. That research could include not only the full range of biological and physical sciences but also linguistics. With 170 separate Indian languages in Brazil, there is a unique opportunity to take the world lead in the field. This has been partially recognized in a special government project on Indian languages, but the level of activity is still far too low.

There are several other simple things the government could do to help researchers. One would be to remove completely the bureaucratic obstacles that slow the import of scientific equipment and reagents. Government officials say that rules must be applied, but that is not what is in dispute — it is appropriate rules that are needed, not those that regard all imported items as the same.

Other problems are not so easy to deal with because they are financial (the level of funding, the freeze on new appointments in the universities) or political (the unwillingness of the government to allow excellence to be rewarded and incompetence punished among the university's tenured staff).

Clearly the research milieu is now changing and there will be increasing emphasis on interaction with industry, and with boosting the dismally low level of research and development within industry. A few researchers are already seeing the future and making their own plans for new links, others might be wise to imitate them. □

ORGANIZATION

Whistle-stop guided tour

This page provides a map, both geographic and institutional, of science in Brazil. As with any map, its principal appeal is to people who dislike travelling without always knowing exactly where they are. Those who are content simply to arrive may turn to the following pages.

Universities. Brazil has more than a million students enrolled at its federal, state and private universities, but there are only eight universities carrying out active research in a wide range of scientific disciplines.

In a category of its own, well ahead of any competitor, is the University of São Paulo (page 368). It produces more than half the doctorates in Brazil. Second comes the University of Campinas (UNICAMP), located 100 kilometres north of São Paulo (page 368). USP and UNICAMP are not supported by the federal government but by São Paulo state which is Brazil's richest, generating 50 per cent of total gross national product (GNP).

Third is the best of the federal universities, the Federal University of Rio de Janeiro (page 358). After Rio, the rankings are open to dispute (and, of course, any university may have particular specialities in which it excels). But most agree that in the next five places will be the Pontifical Catholic University of Rio de Janeiro (remarkable as the only private university with strength in science, page 359), the State University of São Paulo (yet another of São Paulo's state-funded universities, with campuses throughout the state), the Federal University of San Carlos (again located in São Paulo state), the Federal University of Rio Grande do Sul (located in Porto Alegre in Brazil's southernmost state), the University of Brasília, and the Federal University of Minas Gerais (located in the state capital of Belo

Horizonte).

Research Institutes. At one time or another, all of the major research institutes belonged to what everyone still knows as the National Research Council and refers to as "CNPq". (The council's change of name to the National Council for Scientific and Technological Development has proved unpopular, page 373). Eleven research institutes still belong to CNPq; for historical reasons they are concentrated in Rio de Janeiro. Two major research institutes, for space research and amazonian research, are now attached directly to the what most people refer to as the Ministry of Science and Technology. In fact, the name is not correct. In a dramatic political battle, the ministry was absorbed by another ministry early this year but fought its way out again, victorious and independent, as the Special Secretariat for Science and Technology of the President of the Republic (page 371) a ministry except in name. The research council, CNPq, is answerable to this body.

There are many other research institutes — for Meteorology, Metrology, Mining and so on — attached to other ministries but they are less central to *Nature's* interests. Exceptions are the Oswaldo Cruz Foundation of the Ministry of Health, and the EMBRAPA chain of research institutes attached to the Ministry of Agriculture (pages 370 and 363).

Research support. Virtually all support for scientific research comes from the government. The total budget in 1987, the last year for which complete figures are available, came to \$1,640 million, approximately 0.7 per cent of GNP. The following year the budget rose but in January 1989 massive cuts were ordered by the President. As supplementary budgets are still being debated by Congress it is not possible to say accurately how much will be spent this year, although the final figure could be as low as half that granted in 1987. In that year the budget was divided up among ministries as shown in the pie-chart below, with the Ministry of Science and Technology receiving \$470 million (35 per cent of the total). Of that \$470 million, just over half (\$249 million) went to CNPq to support its scholarships, research grants and institutions (page 373). Thirty per cent went to the Agency for Financing Funds and Projects (FINEP), which is part grant agency and part

Major research institutes

CNPq

Rio de Janeiro

National Observatory (p.359)

Museum of Astronomy and Related Sciences

National Laboratory for Astrophysics (p.374)

Brazilian Center for Research in Physics (p.361)

Institute of Pure and Applied Mathematics

National Laboratory of Scientific Computing

Center for Mineral Technology

Belém

Emílio Goeldi Museum (p.363)

Campinas (SP)

National Synchrotron Light Laboratory (p.368)

Brasília

Brazilian Institute for Scientific and

Technical Information

Centre of Science and Technology Policy Studies

Special Secretariat (Ministry of Science and Technology)

São José dos Campos (SP)

National Space Research Institute (p.367)

Manaus

National Institute for Amazonian Research (p.362)

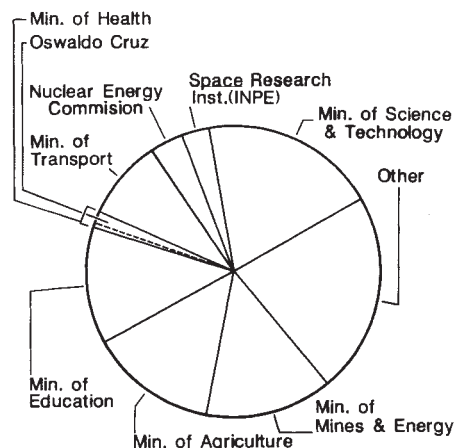
Ministry of Health

Rio de Janeiro

Oswaldo Cruz Foundation (p.370).

commercial bank (page 372), and to two other ministry funds, the National Fund for Scientific and Technological Development (FNDCT) and the Programme of Support for Scientific and Technological Development (PADCT). The latter is particularly important because it is through this fund that the World Bank has made a \$72 million matching loan to Brazil to help in the development of the science base (page 372).

Outside of the Ministry for Science and Technology, support for basic research



The distribution of government research spending by ministry during 1987.

comes from the Ministry of Education's Coordinating Agency for Advanced Training of High Level Personnel (CAPES), which supports graduate training in the universities (page 372). Individual states also have their own funding agencies, São Paulo's FAPESP is the best known and biggest (page 372). □



Island exile for academics

Rio de Janeiro

FOR all its carnival image, Rio on a rainy day is about as exciting as Manchester, and considerably uglier. Rio is more than a few famous beaches and a soccer team; it is a city of 5.7 million people (a million of them living in slums), an industrial centre and a major port. Heading north out of Rio, the Guanabara Bay elevated expressway runs for two grimy miles alongside rows of wharves, cranes and warehouses before winding out of the city through the industrial Zona Norte. There are no tourists out there.

But Rio in the sun is another city altogether — certainly one of the most beautiful in the world — not because of anything man has built but because of its mountains and the sea. The famous Sugarloaf is remarkable only for its symmetry among an infinite variety of mountain forms — some bare rock and some covered with lush forest — spilling down to the sea and a series of beaches, bays and lagoons. There is nothing to add to the clichéd speechlessness of Charles Darwin who wrote, visiting on the *Beagle* in 1832, that the landscape “so completely surpasses in magnificence all that the European has ever beheld in his own country, that he knows not how to express his feelings”.

Rio residents add utility to aesthetics; they like to say that the only reason that major riots do not break out in the slums — the *favelas* — is because of the open beaches and overwhelming beauty of the surroundings. While the idea of beauty as a palliative for poverty may seem romantic, in Rio it is believable.

Rio was long Brazil's capital and tries to claim the honour of being the first city to open a proper university, as against faculties or schools for the study of individual disciplines. But the motive for setting up the university was not entirely scholarly, or so the story goes. Diplomatic necessity was the driving force. The King of Belgium was about to pay a state visit to Rio, and the president found that he had no university to confer a suitable honorary degree upon his distinguished visitor. The University of Brazil was thus created, although it was not to last for long. Some of that university's buildings still remain in an old part of town, with a pleasant view across the bay to the Sugarloaf. Alongside it are two of the research institutes of the

National Research Council (CNPq): the Brazilian Institute of Physics (CBPF) and the National Laboratory of Scientific Computing (LNCC). A third CNPq institute — the Institute of Pure and Applied Mathematics — also occupies a delightful site overlooking the botanical gardens on the lagoon behind Ipanema. And not far from there, also surrounded by greenery and mountains is the Catholic University, PUC.

But Rio's main university, the Federal University of Rio de Janeiro, enjoys no such picturesque setting. Almost the entire university was moved in 1972 to a low, uninhabited island, five miles from downtown Rio, just north of the docks of Guanabara Bay. Many academics claim

city. But it seems unlikely that the island will really begin to feel lived in and friendly before the end of the century. The university has 25,000 students and 3,000 faculty. Overall, it is usually described as the third-best university in Brazil, but its biophysics group is the best and its reputation in biochemistry and engineering, where the university runs a very large and innovative graduate school (COPPE), is very high.

The university faces the now classical problem of other federal universities — growth has stopped. For the past four years a federal freeze has prevented the university hiring any full-time staff (there are ways of hiring people on specific contracts). People returning from post doctoral experience abroad must thus face the dismal prospect that the universities are full of still-young people, much less qualified than themselves, who joined the university a decade or more ago. It will be the next century before the hiring bulge of the 1970s fully works itself out of the university system.

Rio has always had the pick of good graduates so the problems are not so severe here, but Professor Luiz Pinguelli Rosa, director of COPPE, explains that many universities are “dominated by people who are not concerned with research, the old-fashioned teacher, professional and bureaucratic people who just come to class and then go home, and now with political appointees who have nothing to do. But now is a time of transition”. He also

The campus architecture in Rio: not an object lesson.

that the the university was moved for political reasons, so that out on the island the military could be sure the students could foment no trouble.

Other academics claim the story is an exaggeration; they say the island is simply a manifestation of the “gigantism” that comes easily to the totalitarian mind. The island is artificial, created by filling in the land between eight tiny islands, and took almost twenty years of intermittent work to complete. It is two miles long and for the most part, covered in an untidy scrub. At one end stands a vast university hospital, at the other a cluster of university faculties, housed in concrete blocks, but each so far from its neighbour that it seems unlikely that there are many spontaneous meetings between the members of different faculties. Rows of students wait at the bus stops of each faculty or try to hitch a ride back to Rio. There is nowhere to live nearby.

Optimists say that it is too early to judge the project as the island will gradually fill in — new research institutes linked to private and state companies are already being built — and become an academic

points out a certain irony in the present situation. Many academics opposed the military and, in the late 1960s, large numbers were dismissed or even driven into exile. But despite the military's pressure on academic freedom, they expanded the universities and put money into science and technology. Now the government has lost the will to invest for the long term.

Despite the shortages it is possible to keep good groups going. Nelson Maculan, a COPPE professor, has a team of young researchers with postdoctoral experience at universities that read like a listing of the world's best. By early October he had already received 25 visiting professors since the start of the year. But his field is theoretical — operational research — and he needs only a work station and outside contacts to stay at an international level. He despairs of how to bring back good people who have gone abroad for training when he has no jobs to offer them.

Rio is also swept up in the wave of democracy passing through the universities, another reaction to the disappearance of the military. Everywhere students, and particularly support staff who



greatly outnumber faculty, are asking for a bigger say in running the universities and in electing their rectors.

Rio has gone further than most, prompting the Minister for Science and Technology, Décio Leal de Zagottis, to claim that the a "crisis" is imminent that will harm the university for a decade, slowing its attempt to form new links with

CATHOLIC UNIVERSITY

industry. The courts will ultimately have to decide whether, in de Zagottis's words, a "university can be directed by students, workers and professors, all at the same level". But academics at the university say the problems are being exaggerated, and Bio-Rio, a key biotechnology venture on the island (see page 374) says its immediate prospects are good. □

The art of living on the edge

Rio de Janeiro

ONE of the most studied announcements on the noticeboards at the Pontifical Catholic University of Rio de Janeiro (or PUC as everyone calls it) gives the latest prediction of when staff salaries might be paid. PUC is the only private university in Brazil carrying out a substantial scientific research programme, but this remarkable achievement is rewarded irregularly in these difficult times — the salaries for January, for example, did not arrive until April. "It is a very unstable situation", says Professor Humberto Brandi, Dean of the Technical-Scientific Center (CTC) which coordinates the university's scientific programmes. "We have constant problems of money . . . I don't think you can imagine how it works. It's completely crazy."

The private status of PUC makes it particularly vulnerable to economic instability as it does not have the guaranteed income of federal and state universities and must rely instead on student fees, and research grants and contracts.

PUC was set up in 1941 as Brazil's first private university. In 1959, the university bought Brazil's first computer, a gigantic US-manufactured machine that took up two floors and was part mechanical and part electronic. When the university found the computer could not be switched off, they were stimulated to appoint the first full-time staff. The university is still very strong in computer science and telecommunications; IBM is just about to donate a large mainframe computer and the telecommunications research centre (CETUC) receives large contracts from the state telecommunications organization, TELEBRAS.

The university has 8,000 undergraduate students, 3,000 of them in the Scientific-Technical Centre. The emphasis is on technology, rather than basic science. There are 1,000 graduate students in the centre, along with 260 faculty.

The faculty's funds and salaries come mostly from the research council FINEP (see page 372) and increasingly from research contracts (PUC has around 160 of them) from state and private industry. But early in October, Brandi is still waiting for September's salaries to arrive. In April, when they had been without sal-

aries from January, he had been forced to borrow some money intended for investment in projects to pay part of the salary bill, then later repay the money, with interest. As interest rates are currently 1.8–2.0 per cent a day, this is a painful transaction. Asked why he appears so calm about it, he relies, "We have been living on the edge for ten years, we know how to do it".

This time, anyway, he says he knows the money for salaries is there, it just has to be "liberated". "Liberation" is a word that senior Brazilian researchers use often. It means a kind of guerilla-style harassment of government agencies until they deliver the funds that are owed. The budget is decided by the Ministry of Planning and is approved by Congress but the funds are disbursed by the Ministry of Finance. There the officials are under constant pressure from the International Monetary Fund to improve the deficit. One trick is to slow the pace of payment; because the budget is not inflation indexed, every delay means a decline in the value of the amount to be given out. Brandi and other "liberators" must thus spend their time on

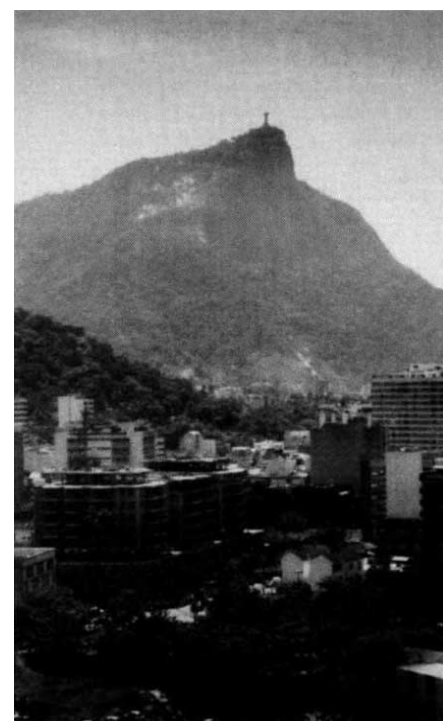
NATIONAL OBSERVATORY

Counting pennies not stars

Rio de Janeiro

It is the middle of October and Ramiro Muniz, vice-director of the National Observatory remarks, without any particular emotion, that the observatory is absolutely broke. "Today we do not have a single cent", he says. All that is left are some funds earmarked to pay the electricity bill.

Muniz hopes that Congress will soon pass a supplementary budget, designed to resurrect funds that melted away under 30 per cent a month inflation. He explains that the observatory operating budget (which at the beginning of the year would have been worth \$800,000) exists in a computerized account in Brasília. It bears no interest, nor is it corrected for inflation. It appears at the beginning of the year, and immediately begins to lose one per cent or so in value every day. The observatory tries to buy as much as it can immediately — but as Muniz says, "you



The famous statue of Christ looks down on the Catholic University.

the telephone, and flying up to Brasília, trying to make the funds move down the pipeline.

"Brazil is a young country and everything is very informal", says Brandi. "It's possible to pick up the telephone and talk to the minister. The ministry has its own priorities which are complicated, and works on pressure." Last month's salary has, at any rate, been liberated from the Ministry of Finance and travelled as far as FINEP. Brandi is now occupied with liberating it from there. □

cannot go hog-wild buying everything in January".

Anyway, the law prevents advance payment of government funds to private companies without special bank guarantees; it's payment on delivery, and by the time a delivery arrives, costs may have doubled, or tripled. . . .

Do-it-yourself becomes the only way to save money. The observatory has taken to mending its own cars and using its technical expertise to repair building — the workshop apparently produced a fine set of aluminium doors. Muniz clearly likes improvising. He is a physicist (and spent five years at Berkeley in the fifties) but most enjoys the times when he can use his knowledge of paramagnetic resonance in unconventional settings: verifying a colleague's hypothesis that a particular drug acts on trypanosomes through the release of free radicals, and finding that the magnetic signature left in rock cores can

be used to recover temperature history. Now, necessity has made him an expert on economics and accounting. He expects no quick solutions to Brazil's problems. "The politicians are as lost as a blind man in a shoot out", he says, "they have no idea which way to run".

The observatory also has internal problems which are essentially a reflection of its origins. The observatory began in the early nineteenth century as a practical place: it made astronomical observations of use in navigation, recorded magnetic data and corrected compasses, and kept the hour, supplying a time signal at noon for ships in Rio de Janeiro harbour to set their chronometers.

The observatory (and its old telescopes) now stands on a little mango-tree covered hill in the São Cristóvão region of Rio; its original hill, which was closer to the sea, was quarried away sixty years ago and used to extend Rio's waterfront. Today, the observatory is unhappily split amongst the descendants of its original functions. There are astronomers (who now see themselves as basic researchers), geophysicists (who have forged links with industry), and engineers who still keep the hour (and are looked down on by the first two groups). Muniz tries to prevent argument



The National Observatory's roots lie in the early 1800s.

over the groups' respective merits, and impishly quotes Lewis Carroll on the reality behind words, "The question is, said Humpty-Dumpty, 'which is to be master — that's all' ". Muniz's view is that "it is less important who is in charge inside than for the observatory to be good enough to get funds from outside".

Brazil has maybe 100 working astronomers (the astronomy society has 300 members), the great majority of them either at the National Observatory or at the University of São Paulo. The main observing facility is the National Astrophysics Laboratory, associated with the National Observatory and located at

Brasópolis in the mountains some 300 km from Rio. Groups from several universities use the facility, which is equipped with a 1.6 m reflector and 60 cm refractors. Brazil is not the best place for optical astronomy; there is too much cloud and not enough high mountain.

If support is forthcoming it is likely that Brazil will use a facility in the Chilean Andes, probably in collaboration with a US university group. Some observations are already made in Chile and Argentina, as well as in South Africa. Projects at the observatory include the South American part of the Harvard-Smithsonian redshift survey, a stellar astrophysics group that studies the polarized emissions from stars, and a wholly Brazilian project on T-Tauri, young stars.

Elsewhere in Brazil there is a 30-metre SBPC

radiotelescope (now at INPE, at one time it belonged to the observatory) but it is already old. A consortium from the University of São Paulo and INPE and some other groups — the first time a consortium approach has been tried out in Brazil — is now attempting to find funds for a much larger radio array.

The observatory's geophysicists are responsible for Brazil's magnetic and gravimetric surveys. They also do seismic work with Petrobras, the state oil company. North-east Brazil is seismically active although no shock bigger than Richter 5 has been recorded. Much larger earthquakes are believed to have passed unrecorded in the Amazon, and there are plans to set up a seismic station at Manaus — when and if the money becomes available. □

Science as continuous struggle

Rio de Janeiro

ENNIO Candotti is a name known to virtually all Brazilian scientists, which is some measure of the continuing importance of the Brazilian Society for the Progress of Science (SBPC), of which he became president this year. SBPC is the exact equivalent of the AAAS in the United States, it organizes seminars and holds an annual meeting (which as many as 10,000 people attend), it lobbies the government on behalf of scientists, and publishes a scientific magazine.

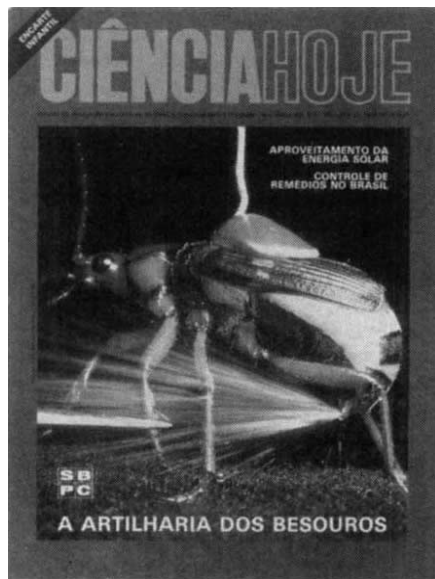
Candotti is a busy man, constantly in demand. He is not without his critics, who fear he might not be big enough to force the government to provide science with the support it needs, or suspect he has become too much of a journalist after having spent seven years as editor of SBPC's monthly magazine, *Ciência Hoje*.

By training, Candotti is a theoretical physicist. Over a lunch, taken late in the kitchen of the Oswaldo Cruz Institute (the official lunch in the dining room had ended an hour earlier and all the guests had gone home, but a series of interruptions, and people queuing to talk to him, had so far prevented Candotti from actually eating anything), he characterized, somewhat melancholically, the state of Brazilian science as "continuous struggle".

"For each hour of investigation, there are ten hours of fight to get the money, to get equipment, to solve the problems of stability for your laboratory. Every person is divided between two jobs, one is the old job we used to have, the other is the fight to maintain the old job. . . . We dream of a stable budget, to have equipment, to have libraries, but we don't get them". The result is that Brazilian scientists have become very ingenious. "If you gave them half of what an average US scientist has, they would be able to achieve three or four

times as much", Candotti says, half-seriously.

"Political interests are always present in scientific life," says Candotti. "Why? Because the whole Brazilian problem is political. Fifty million people are under-educated and without health care, 30 million are without basic instruction, 60 million earn less than \$60 a month. At the same time we have a very powerful industry and economic growth. It is these contradictions that must be solved to give stability. We need 100,000 researchers, we have 30,000, we don't have people to do the most simple things. But the government does not have the stability itself to provide continuity to a programme for



SBPC's monthly magazine, *Ciência Hoje* (Science Today), has a circulation of 70,000 and is "written by scientists, rewritten by journalists", according to SBPC director, Ennio Candotti. It is modelled on *Scientific American* and *La Recherche*.

four or five years . . . 30 per cent a month inflation destroys all plans and budgets."

Inflation is something Candotti had direct experience with as editor of *Ciência Hoje*. "I buy so many tonnes of paper, but one month to the next costs can grow by 50 per cent; and by the time you receive the magazine subscriptions they are worth nothing, so how can you do the next issue?" Candotti's job now is part lobbyist for the scientific community. SPBC is not the only lobby group; the 60 scientific

Folha de São Paulo



SBPC's Candotti: The table is bare.

societies of Brazil also have a nine-person commission to lobby Congress and the ministries, and other groups represent individual universities and research institutes. Although some argue that the appearance of these new groups means that the SBPC is not as strong as it was, a better explanation might be that society is maturing and SPBC no longer need be the only opposition group, now that military ruler has ended.

Candotti welcomes the appearance of other groups. "In the last year or so, society is becoming more organized, although there are no true political parties, there are more pressure groups. In the recent discussion of the new constitution, many groups were organized. Perhaps this will change the next government, it will not be completely free to govern by authoritarian methods, but will have to work with society. Congress has a new and important role".

So far, he says, there has been too much arbitrary political influence over scientific programmes. "If the minister changes, then the whole scientific enterprise can change 180 degrees. Elsewhere they are more pragmatic, here things can change from one direction to the opposite . . . all programmes, all priorities can be changed in a few months."

While Brazil may be more unstable, the job of lobbying the government is little different from elsewhere. Politicians give you attention "according to your political force, who you represent and who you can mobilize", says Candotti. □

Physics in times good and bad

Rio de Janeiro

PROFESSOR José Leite Lopes' office in the Brazilian Center for Research in Physics (Centro Brasileiro de Pesquisas Físicas, CBPF), where he is director, is decorated with photographs of Wolfgang Pauli (who he knew when he was in Princeton), and Albert Einstein (taken on a chance encounter); the corridors nearby bear giant pictures of Richard Feynman and Brazilian colleagues, taken on one of his several stays at CBPF. The treasure among them is one of Feynman dressed as Mephistophiles for the 1952 Rio carnival.

Those early days were exciting — the beginning of physics in Brazil. In 1947 Cesar M. G. Lattes, then just 21 years old, had taken a set of exposed photographic plates from Chacaltaya, the cosmic ray observatory high in the Bolivian Andes, to C. F. Powell's laboratory in Bristol, England. The result was the discovery of the pions and their decay which was published by Lattes and Powell in *Nature* in 1947 and won Powell the 1950 Nobel prize. Leite Lopes, who was then in his mid-twenties, published his own work on meson decay and the theory of nuclear forces, also in *Nature*, in 1948. The next year, these two young scientists (Lattes, fresh from Berkeley, where he had won further fame and had appeared on the cover of *Life* magazine) set up the institute where Leite Lopes is now director. "The university", says Leite Lopes, "did not understand research, so we set up our own organization, beginning as a private institute and helped by the prestige of Lattes". A few

months later Feynmann came for a couple of months, he was back in 1951 for his sabbatical year, and again afterwards. There were other famous visitors too — J. Robert Oppenheimer, Eugene Wigner and Chen Ning Yang among them. But this lively period came to an end with the military coup in 1964. After being briefly detained, Leite Lopes moved to Paris, but returned to Rio in 1967. Two years later he was expelled by the military government. "I was clearly a progressive," he says. He moved first to the United States and then to the University of Strasbourg, where he remained in exile for 15 years, "working hard and living well — a gift from the generals". In 1985, with military rule ending, he was invited back, to be director of CBPF for the second time. In the meantime the institute had become one of the 11 CNPq research institutes.

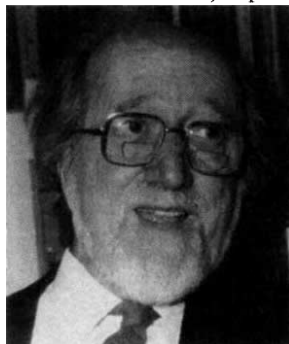
Leite Lopes' "progressive" leanings still show; although the problems of the institute (simply put: "we have no money . . .

if they take anything more we can only go to the beach") are foremost in his mind, he bunches his fists with anger when he talks about the state of Brazilian society, the ten million children living in *favelas*, and the breakdown of the public education system that had provided for poor people.

Back at the institute, Leite Lopes' main task has been to try to build the experimental physics that is largely lacking from the Brazilian scene, except at the University of São Paulo. The institute's own linear accelerators are 30 years old and no longer work. One success was the designation of the institute as the National Laboratory of Nuclear Fusion. But the plan, to study muon-catalysed fusion, exists only on paper.

Despite the lack of sophisticated equipment, several groups at CBPF look after themselves very well. Some have worked out ways to work with scientists abroad, others have adapted their science to the equipment available in Brazil. The institute has 93 researchers and more than a hundred graduate students.

The institute still has several projects going on at the Chacaltaya cosmic-ray observatory, now in cooperation with the University of Tokyo's Institute for Cosmic Ray Research. But the jewel in the crown is the group of 20 high-energy physicists, led by Joao Anjos, which works closely with Fermilab in the United States; the group even had President José Sarney pay Fermilab an official visit in September. The group has built (at a surprisingly low cost) a parallel multiprocessor, a twin of



Leite Lopes: returned exile.

the ACP machine at Fermilab, enabling all data gathered at Fermilab to be analysed back in Rio.

A group studying hyperfine interactions has taken the option of tailoring its science to available equipment, rather than gathering data abroad. CBPF pioneered research on the Mössbauer effect in Brazil, and has carried out the most complete Mössbauer studies

of the new high-temperature superconducting materials.

Leite Lopes is thinking of returning to France to an emeritus professorship at Strasbourg and leaving the problems of Brazil to someone younger. His critics (of which he has rather few) do not mention his age, but that he is too big a figure for today's Brazil. The same comment is heard of others from his period (including University of São Paulo rector, José Goldemberg), particularly of those who spent time in exile. As one researcher put it, "only giants could still do good science through all those times". □

Fighting for the Amazon

Belém

ON the waterfront in Belém, flat-bottomed river boats, hammocks swinging between their open decks, bring all the curiosities of the Amazon to market. There are huge fish and unfamiliar fruits. A row of stalls sells medicines and charms — herbal oils, lizard skulls, alligators' teeth, dolphins' penises, turtle shells, and blue-and-white banded snakes, with milky, opal eyes, suspended in jars of fluid. Close by, a little shop sells the curious paraphernalia of the African-derived Umbanda religion, which includes among its pantheon St George (complete in shield and armour, slaying the dragon), Iemanjá, white-robed African goddess of the sea, and Pai João, a cigar-smoking black slave.

Even the ice creams are Amazonian; alongside the familiar tropical tastes of pineapple, guava and mango are açaí, bacuri, caju, cupuaçu, muruci, graviola and genipapo*. Several of these fruits are collected from the forest and not yet grown in cultivation. Even so, it is just a matter of time before they reach New York: the agricultural research station in Belém has worked out how to spin-dry and powder some of the fruits for long-term storage.

Belém is just south of the Equator, 2,000 miles north of Rio de Janeiro and 80

miles from the open sea. It is the capital of Pará, a state almost six times the size of Great Britain. Although an exotic place for the visitor, Belém is also the home of three of the most important scientific institutions within Amazônia: the Emilio Goeldi Museum, which is the region's oldest scientific institution, the Federal University of Pará, and the Centre for Agricultural Research in the Humid Tropics (Centro de Pesquisa Agropecuária do Tropicó Úmido, CPATU). The heads of all three institutions have much on their minds — the parlous state of their finances, politics, their research plans and, not surprisingly, the emergence of the Amazon as a world issue. On this latter issue they all agree that rational, sustainable development of the Amazon is the only way forward (not the irrational development of some of Brazil's giant projects, nor the rational non-development of diehard conservationists), but they disagree on other points, particularly on how well Brazil has done so far and on the role foreigners might play. □

*These fruits are *Euterpe oleracea*, *Platonia insignis*, *Anacardium occidentale*, *Theobroma grandiflorum*, *Byrsonima crassifolia*, *Annona muricata* and *Genipa americana*. Some are found only in the Amazon region, others are cultivated throughout the tropics.

An Amazon University for Amazônia

NILSON Pinto Oliveira is rector of the Federal University of Pará, which occupies a campus down by the river on the outskirts of Belém. The university would not come on anyone's list of the best universities in Brazil, but it is the best in the Amazon — and not just the Brazilian Amazon, but the whole region.

Oliveira knows the potential of the region and the reality of his resources. The university has 1,500 faculty but just 135 with doctoral degrees — one trained researcher for every 36,000 square kilometres of the Amazon. But as Oliveira says, there is no time to stop, "the Amazon is not simply a place to admire, it's a country where people live and will continue to live. The people are poor, they need money for the next day."

Oliveira has been rector for three months and clearly chafes at the many restrictions under which he has to operate. His first priority is to raise the level of training of the university staff and help improve Amazon educational standards. "We cannot talk about the protection of the Amazon region if we do not give to the Amazon people the possibility of learning and understanding their problems", he says. The university, along with other institutions in the region, is running a scheme to retrain school teachers. At sites

throughout the Amazon, some 2,000 schoolteachers come to twice-yearly three-month courses. Raising university standards is exceedingly difficult. Most of the faculty, perhaps 90 per cent, do no research but simply attend the university to teach. A considerable number bend the rules and hold other jobs elsewhere. The four-year-old freeze on appointments prevents new staff being brought into the university. And central control by the Ministry of Education prevents change.

If Oliveira had his way, he would prefer 500 well-trained researchers than the 1,500 staff he has now. He sees lack of flexibility as an even bigger problem than inflation or shortage of money. "All universities are considered to be the same by the ministry", he says, "But the priorities here are completely different from the University of Rio de Janeiro. I cannot create posts in the area they are needed."

The university hopes to consolidate its resources in environmental research in a Centre for Ecological Science. Foreign help is needed to train people.

But Oliveira, like others in Brazil, is convinced that foreign help must be true co-operation which will build up expertise in the Amazon.

Perhaps 70 to 80 people at the university are involved in research. Geoscience is one active area and is involved in a collaborative project funded by the US NSF to study the Amazon shelf off Marajó Island. Another active basic research group, in primate genetics and evolution, is run by the husband-and-wife team of Horacio and Maria Cruz Schneider.

Their research has received a bonus from the construction of some of the



Oliveira: making the best of things.

Amazon's controversial hydroelectric dams. When the reservoirs are flooded, power companies are under obligation to catch large animals trapped by the rising waters and release them elsewhere. The Schneiders collect blood samples from captured monkeys. Study of blood groups, serum proteins and chromosome types enables questions of New World primate speciation and evolution to be tackled.

The Schneiders' efforts are quite heroic. The laboratory roof sometimes leaks, the water and electricity supply are unreliable, and 90 per cent of their electrophoretic reagents have to be imported with all the delay and bureaucratic uncertainties that involves. The university is too poor to afford journal subscriptions. Instead, the primate group has banded together to pay for airmail delivery of *Current Contents*. They then write for reprints to the authors of relevant work. The team has great spirit; all 15 primate researchers are young and are committed to refusing any side jobs. They are also building their international links; members of the group are slowly being 'strategically placed' on fellowships at universities in the United States, Japan and Europe. □



The Federal University of Pará's riverside campus.

Museum ups and down

GUILHERME de La Penha is director of the Museu Paraense Emilio Goeldi, to give it its full title. The museum is old, established in 1866, and looks as though it has emerged from the pages of a Persian miniature painting. In a densely wooded, walled botanical garden are set pools and aviaries full of the birds and beasts of the Amazon: manatees, jaguars, giant river otters, cranes and toucans — 600 species in all. Almost half a million visitors come here each year. Its scientific heyday was early in the century when the Swiss zoologist Emil August Goeldi was director and began its collections of biological, ethnographic, archaeological and geological specimens.

The past few years have seen a revival at the museum, for which de La Penha can claim some credit. The museum is quite small, with around 100 scientists, and has been able to adapt to changing circumstances more quickly than INPA, the much larger Amazon research institute in Manaus.

De La Penha says some 30 or so of the staff are now producing "good science", and have won back an international reputation for the museum. Both the Ford Foundation and the MacArthur Foundation are supporting research there and more than 40 people have been sent away for postdoctoral training. The hope is that their gradual return will take the museum to a much higher level. The museum's rising fortunes were celebrated in October with the grant of 30,000 hectares of natural forest at the junctions of the Xingu and Amazon rivers. A new field research station is to be equipped, most probably with help stemming from an environmental memorandum of understanding signed between Great Britain and Brazil earlier this year. The agreement was a nice coup for the British Embassy's diplomats, having been signed at the height of Brazil's annoyance with foreign 'interference' in its Amazonian policy.

But that is the good news. The bad news is that the museum is very hard hit by this year's cuts in funding for science. De La Penha appears under siege. "We cannot buy equipment, we cannot buy books. Everything that is important and imported has become impossible for us", he says. Inflation has also taken its toll. "The value of salaries has fallen so much that most scientists are living on their savings", he says. "Now we're always worried about how we will reach the end of the month. You want some peace to publish manuscripts, but you go to bed wondering how can you survive."

Even in Belém, Brazil's political turmoil is felt. The museum is one of CNPq's institutes and CNPq has come out against the idea that all members of an institute —



In the depths of Goeldi's forested zoological garden.

including the large numbers of support staff — should have an equal vote in deciding how it is run and who runs it. But depending on the presidential election, 'democracy' may triumph. De La Penha is disgusted with the intrusion of politics into science. "Incompetents are reaching high positions in the universities and institutes", he says, "although there is meant to be evaluation by peers, peers are afraid the person they evaluate may prove to belong to the right political party. If you're popular you get elected, it doesn't matter if you're competent, you must be popular with students, clerks. . .".

De La Penha is not a stranger to political fights. He left the country in 1983 after quarrels over planning for research in *Amazônia*. He had been vice-president of CNPq and organized a meeting in Santarém to discuss a plan for development of the Amazon. "Even today it is the only written plan for research in the Amazon", he says, "then we tried to put on pressure to create a tropical research programme, we also criticized the Interior Ministry, showing that many of its project funds just supported the infrastructure of institutes

and were not used for research." These and other activities got de La Penha into hot water, and he accepted 'exile' to a job at the Organization of American States in Washington, DC. But two years ago he returned to Belém. His aim now is to establish a small, very strong scientific institute and not to interfere with the Amazon development agencies. "If it is strong it will be heard", he says, "and will provide hints, not solutions, of how to develop the Amazon on a rational basis. The Amazon is a complicated set of limited regions that need to be looked after differently, you cannot have a theory of the whole Amazon."

Foreign collaboration is bound to play a big part in the study of the Amazon, says de La Penha. He laughs at politicians who "every time you talk about foreign co-operation in the Amazon, say it's a foreign trick". "We have to have added cooperation. Most of the research on the Amazon is still published by US and UK scientists. They have better continuity and can carry on a project for years. There are segments of Brazilian science who say, leave the Amazon to Brazilian scientists. But that would be one scientist for how many thousand square kilometres? Brazilian scientists must enter the international community. You cannot learn if you attend only regional conferences."

But foreigners who have jumped on the rain-forest bandwagon are another matter. He wishes that some of the 'Save the Rain Forest' groups in the United States would spend their money in Brazil where it is urgently needed, instead of on "parties and lobbying" in Washington.

The museum is desperate for the most basic equipment. De La Penha managed to buy up one truck load of glass containers that were not good enough for the Smithsonian Institution's needs. He hopes to get another, and dreams of good microscopes and, even more, of personal computers. □

Short shrift for foreign do-gooders

ASKING CPATU chief Emeleocípio Botelho de Andrade about his views on foreign interest in conserving the Amazon provokes an explosion. There is "nothing but emotion, disinformation and economic interest" behind international criticism of Brazil's conduct, he says, "foreign powers are afraid that Brazil will develop as quickly as Japan or West Germany".

CPATU's job is to help speed the agricultural development of Brazil's humid tropical regions, so Botelho de Andrade's annoyance over foreign criticism is understandable. The institute is one of the 27 research stations — each responsible for a particular product or agricultural zone — belonging to EMBRAPA, the agriculture ministry's research organization.

Energy policy is a prime example of foreign perfidy, in Botelho de Andrade's

view. West Germany and the United States have taken \$8,000 million from Brazil for the development of nuclear power plants with almost no return on the investment, he says. But when Brazil wants to develop hydroelectric power, the cheap and logical energy source for Brazil, the same nations stir up problems over the Indian peoples who live at dam sites and try to block Brazil's development.

He claims the foreign view of the agricultural capabilities of the Amazon is too pessimistic. The Amazon's soil and vegetation types were mapped in a huge radar project (RADAM) a decade ago. Much of *Rondônia*, where forest burning has provoked greatest outrage, has fertile soil capable of sustaining agriculture, he says. Overall, he estimates that 8 per cent of the Amazon will prove to be highly fertile, an

area of 40 million hectares that could double Brazil's agricultural production.

Included in that area is the 19 million hectares of river margins, the várzea, whose vast agricultural potential is just beginning to be recognized. The river margins flood each year and the soil is replenished by river silt, just as in the Nile. The várzea could be used for growing rice, for example, although no one has found a way successfully to mechanize large-scale production. CPATU's 500 staff, 90 of them researchers, are responsible for evaluating natural resources in the humid

INPA

tropics, selecting and maintaining 'banks' of species suitable for agriculture and developing agricultural production systems. Information is transferred to farmers through state agricultural extension services. Cocoa, rubber, dendé palm, pepper, brazil nuts and buffalo are among the resources studied, as are the many varieties of tropical fruit. CPATU has succeeded in spray drying açaí, making it available as an ice-cream flavour even out of season, as well as guaraná, a coffee-like bean used in Brazil's answer to Coca Cola. □

Essential: a passion for research

Manaus

ON 12 October, Manaus's local daily newspaper ran a large front-page photograph of a bank robber. He lies dead, in a pool of blood, his body mostly covered by sheets of newspaper. A crowd of young men and boys stands around. The atmosphere is that of the famous crime pictures taken by Weegee in New York in the 1940s, right down to the blank, detached stares of the crowd. The difference is that few of the onlookers wear anything more than a pair of baggy shorts.

Manaus, the capital of Amazonas, is civilization itself compared to the gold-rush settlements of Rondônia and Roraima, but it still has the spirit of the frontier. To go any distance out of town you must know how to wrestle a four-wheel drive pickup along dirt roads, or be able to pilot a boat on the Rio Negro. The city has electronics factories and even a century-old rubber-boom opera house. But just outside its limits, the forest begins and continues for a thousand kilometres in every direction. The researchers who live here are an unusual and independent breed. All are in some way connected to INPA, the National Institute for Amazonian Research, which occupies a sprawling pair of campuses on the edge of town. The institute has a high international reputation but is unfortunately almost bankrupt. In October, it had passed the stage where there was no money to buy paper and toner for its photocopiers and it had just run out of toilet paper.

INPA researchers seemed not much dismayed by this latest development; in fact, there is more enthusiasm in an impoverished INPA laboratory than amid the riches of São Paulo. Why? Perhaps because, as long-term INPA researcher Barbara Robertson, put it: "Researchers here have a passion. No one would work here for the money, it's love". Or

as another Brazilian researcher said: "If you're a biologist, or a hydrologist, where else would you want to be at the end of the twentieth century? The Amazon is the last frontier."

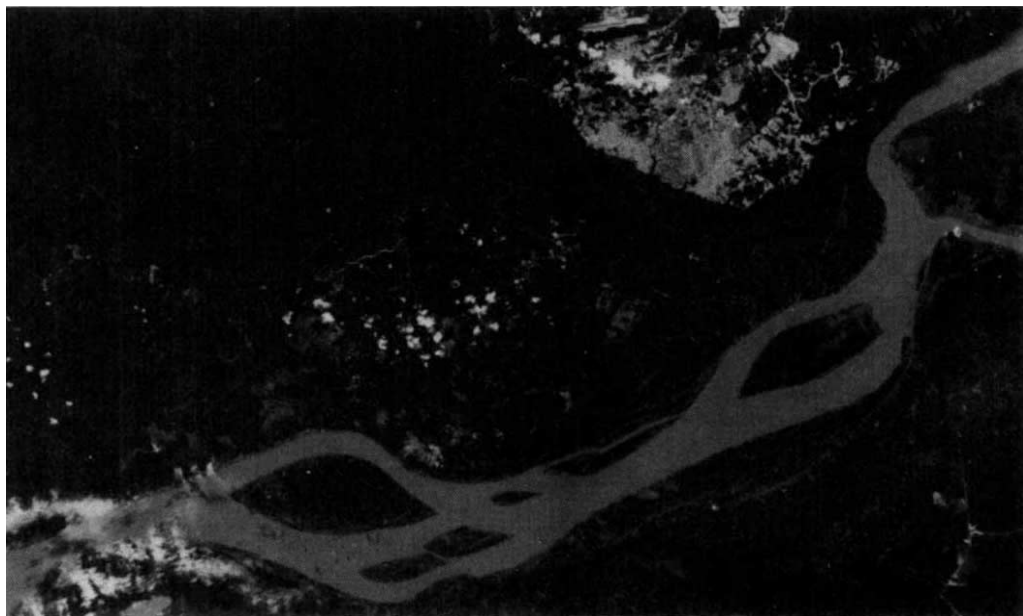
Optimists at INPA say of the financial condition that it is just on a plateau, but the majority view is that it is in decline — despite the rising world interest in Amazônia. Old-timers remember the expansion of the 1970s under director Warwick Kerr, who could be heard constantly yelling at Brasília over the telephone, "putting his job on the line every 15 minutes", as one researcher put it.

The present director, Herbert Schubart, is well known for his calmness amid storm. He points out that INPA has been in very bad shape before but has always managed to win supporters. At the end of the 1960s, some in government wanted to close down the institute, but in the end it was strengthened and moved to a new campus. Now it has 280 researchers and 1,000 employees. The realities are quite different from Kerr's time, he says. In the

1970s, Brazil was in the midst of its 'economic miracle' and INPA grew from 40 to 140 staff in just four years. Political turmoil has proved INPA's undoing. INPA at first benefited when it switched to direct control by the Ministry of Science and Technology, rather than the National Research Council, CNPq. But when this year the ministry was abolished, then reconstituted (see page 373), INPA's budget was lost in the confusion. Not until June was there progress in deciding a figure.

A unique feature of INPA is that close to one fifth of its researchers are foreign, a measure of the international interest in the Amazon. Many are long-term residents and have become thoroughly Brazilian, speaking Portuguese fluently. But their position is sometimes slightly ambiguous. Several have helped to channel money from abroad into big collaborative projects that could never have been funded with Brazilian money alone. One of the more successful is run by Bruce Forsberg, who arrived in Brazil on a post-doctoral fellowship eight years ago and yielded to the Amazon magic. Backed by a group at the University of Washington in Seattle, he has helped arrange a series of research cruises studying the biogeochemistry of the Amazon. With the vast flow of the river (much the greatest on the planet), its seasonal interaction with the flood plain and its effect on the Atlantic, there are several lifetimes of good research projects to do. Unexpectedly, the Amazon flood plain has turned out to be a very important source of methane (a greenhouse gas) as decomposition follows the yearly flood.

US funds backing the project come to a significant percentage of INPA's budget, but the work is totally collaborative, with more Brazilian than US names on their



In this satellite view of the region around Manaus, the black waters of the Rio Negro join the paler Rio Solimões, flowing alongside each other for many miles (see opposite).

papers. This year, for the first time, the group will pick up substantial funding inside the country, from the Bank of Brazil. The funds will provide for a project on mercury pollution of rivers from gold-mining operations by unlicensed prospectors. A preliminary study on the River Madeira in Rondônia has already found dangerous levels of mercury in three locally eaten kinds of fish.

Like several other foreigners at INPA, Forsberg has been able to become an official employee. He has a tiny office, about the size of an American cupboard, in the centre of which he perches on a curious home-made stool, apparently of his own invention. Phillip Fearnside is another foreign staff member and the biggest star internationally. His work on human settlement in the Amazon and his critiques of the giant development projects — the TransAmazon highway, the Balbina dam, Ludwig's Jari colony and the Carajás smelting project — have turned him from lone academic to international spokesman. To his credit, Fearnside's studies have won tremendous respect inside Brazil; despite his criticism of many government projects, he is not regarded as a mere interfering foreigner. Fearnside, a very tall and very thin American, with a moustache that would be the envy of a walrus, is still a quiet academic. But he seems to enjoy the fact that complete strangers sometimes recognize him from his television appearances, even in the wilds of Rondônia.

That foreigners can become employees, and that they can apply for Brazilian research grants, is evidence of open-mindedness on the part of the Brazilian scientific community. But that foreign researchers sometimes obtain both Brazilian and foreign funds can also provoke envy. And regrettably, some insensitive foreigners have come here like colonists of old, completing their projects and departing, leaving nothing behind for their hosts. INPA is considering whether it should begin charging overheads to visitors.

Brasília too is sometimes heavy-handed. Foreign collaborations do require a complete, matching set of Brazilian and foreign partners, but application of the rules is often cumbersome and has slowed good projects.

Foreigners are luckily immune from the institute's politics, of which there is far too much. Among its most senior people are some with poor scientific credentials whose ascendancy owes much to good political connections. But that perhaps just adds another dimension to an institute whose main quality is enormous diversity in approach, quality and style. There are researchers — both Brazilian and foreign — with an international reputation and others who rarely



Boat people. The well equipped floating laboratory, crewed by scientists from the Max-Planck-Institute (see below).

publish, even in an obscure local journal. There are researchers such as Ilse Walker, who is as happy arguing about population biology or Eigen's hypercycle theory (or anything else) as she is on the river collecting zooplankton, her speciality. Others appear not to have read an international journal in a decade. And there is the mix of people pursuing quite practical problems, such as the suitability of Amazon timber for building, and those whose interests are in basic

research, modelling energy flows in ecosystems, for example.

Such heterogeneity may be both stimulating and reveal a weakness: in a visit of only a couple of days, it is possible to uncover researchers who have never heard of one another, even though their work has considerable common ground. That surely suggests a loss of common purpose. Perhaps it is time someone began to organize some institutional seminars. □

Almost too many species to count

CLOSE to Manaus, the Rio Negro is several miles wide. Out in a boat the eye is drawn not to the river, nor to the green line of the forest, but to the sky, a vast sky like that over the ocean or the prairies. Over the day, a series of giant clouds form, build into fantastic storm shapes and then resolve themselves in short, torrential rain showers (see following page). Here the Amazon is flat and the views distant; the river is descending towards the sea at a mere metre every 100 km.

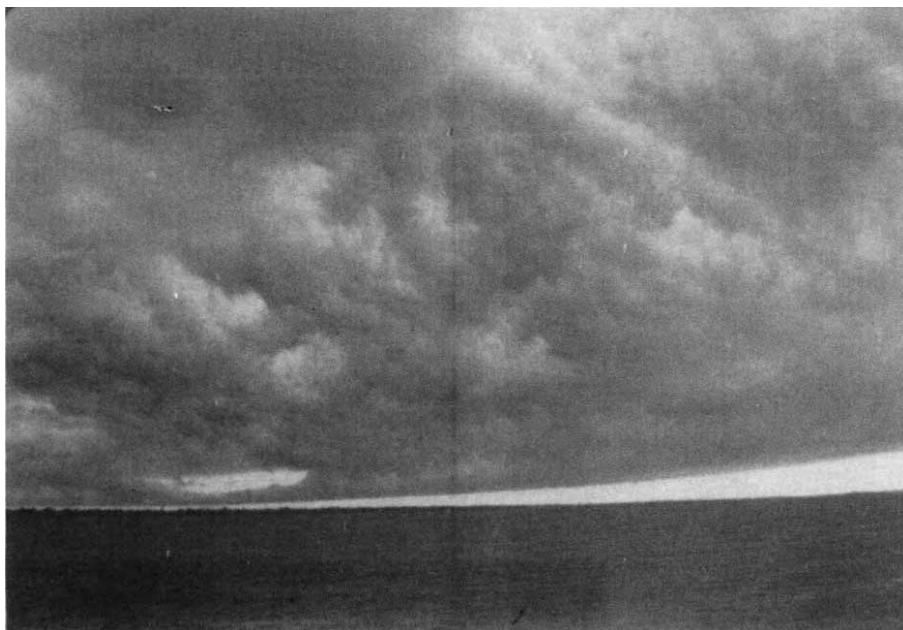
Upstream is the Anavilhanas, the world's largest freshwater archipelago, where the Rio Negro divides into a myriad channels. Downstream a few miles the black waters of the Rio Negro join the pale brown waters of the Rio Solimões. The two flow together for miles, mixing only slowly to form the Amazon. Freshwater dolphin live here and follow passing boats; one species is a curious shade of shocking pink.

A little distance up the Solimões, anchored in a channel away from the main river, is a floating laboratory run by an INPA-based group from the Max-Planck-Institute for Limnology at Plön. It is a meditative spot, the silence broken only by the cry of birds (some of which have a

startling resemblance to the cormorants and kingfishers of Europe). Small boats pass by now and again in a larger channel, destination unknown. One researcher speculates out loud if it might be nice just to buy a boat and set off wandering up the Amazon.

There is a host of interesting problems that can be studied from here, including the hydrology and chemistry of the river itself. But many researchers have turned to the interaction between the river and the land, which is both more visible and more complex than in temperate zones.

The banks are muddy, not because there are tides, but because the river may fall as much as 15 metres between the high season in June and the low in November. Each year, a cycle is repeated in which shore-dwelling plants and animals are first inundated and then left high and dry, and river fish move in beneath the flooded forest and then out again with the retreating flood waters. The result is a unique environment in which plants and animals must adapt to both land and water phases. There are fish that live exclusively on the fruit falling from the forests, others that breathe air to get them through the flood phase when organic matter is decompos-



Storm clouds gather daily over the river and disperse forcefully.

ing in such quantities that the river waters are starved of oxygen.

As in other Amazonian ecosystems, there are huge numbers of new species waiting to be discovered. But researchers

quickly pass from being astonished to being irritated by their numbers: there are so many to be named and so few people to do it that the business of classification simply gets in the way. □

Seeing stars in the rain forest

To reach the site of the INPA–World Wildlife Fund project on the ‘Biological Dynamics of Forest Fragments’, you go out of Manaus, past the ‘Noah’s Ark Steakhouse’ with its culinary warning to the varied fauna of Amazônia, and turn off the paved road at a sign reading “Boa Vista, 760 kilometres”. From there, a red, packed-earth road runs straight through the forest, all the way to the capital of Roraima State and on to the Venezuelan border.

The road was bulldozed through the forest in the late 1970s. The terrain undulates lightly; in the dips, the bulldozed earth acts like a small dam and either side of the road large pools of stagnant water collect, full of dead and dying trees. On the rises, areas cleared for cattle pasture can be glimpsed occasionally through the trees. An hour and a half out of Manaus a rough track leads, in 41 km, to the research site. It takes experienced drivers to get through; muddy slopes and swampy hollows have to be rushed in a series of high-speed slides. Frequent stops have to be made to clear away fallen trees.

At the end of the ride, Camp 41 is luxury. It is just a few low, open huts in a forest clearing but it has a picturesque forest pool nearby where one can swim along with the fish. “Hollywood could not have done better”, was the comment of a group of film stars who visited it recently. With the interest in rain-forest preservation running high in the United States, visitors to the camp include actors Tom

Cruise and Robert Redford and singer Olivia Newton-John.

The forest fragments project is one of the best known in the Amazon. The forest at Camp 41 can be thought of as a ‘control’ forest; it is uncut and unspoilt but a marked grid of trails spreading over 10,000 hectares gives easy access to all parts of it. Nearby, cattle ranches encroach on the forest. Amid them square blocks of forest of different sizes — 1, 10 and 100



Forest camp: cookhouse and a place to sling a hammock.

hectares — have been left standing like islands in the pasture.

The project, conceived by Thomas Lovejoy of the Smithsonian Institution ten years ago, is intended to see how isolated forest fragments change. The project has been described as an attempt to find the smallest area that would support a full complement of species, an idea embodied in the project’s old title of “The Minimum Critical Size of Ecosystems Project”. Roger Hutchings, the field director, says that the original idea was too simple, the question has instead become “What conditions are needed for specific groups of animals to survive?” Every species reacts differently to isolation in its own complex way.

Simple conclusions are in any case unlikely, given the heterogeneity of the Amazon. Hutchings is a butterfly specialist. He has found 210 species of butterfly in three one-hectare forest islands. But only 25 per cent of the butterflies are found in all three areas. Why the variation? No one knows for certain, but it may be simply because the rain forest contains huge numbers of species, spread very thinly. Any one small patch will always contain a different rain forest.

These difficulties and the enormous number of species — many of the insects at the study site have never been classified — mean that the project has no hope of being comprehensive, or of following the fate of even a fraction of the species in the forest islands. That is despite some massive efforts: 65,000 trees are marked and are periodically examined, and 40,000 birds have been ringed and released. Much of the project is essentially opportunistic, relying on the interest and availability of researchers; birds, butterflies and frogs are the favourites.

Even to the casual observer, the natural forest and the forest fragments feel different. The natural forest is a quiet and twilight place. Closely spaced trees of an enormous variety of sizes grow straight up to the canopy, 35–40 metres above, without branching. Beneath, there is really no understorey, just scattered low trunkless palms, and a covering of moist, decomposing, leaf litter in dull reds, blacks and browns. There is little sign of animal life except for the ceaseless activity of ants and termites on the ground and the occasional fruit-eating monkey in the trees above.

In a 100-hectare fragment, the vegetation is denser — more like a Hollywood jungle — because light penetrates through the side of the plot. Only towards the centre does the forest appear unchanged. Smaller fragments of 1 or 10 hectares are much more dramatically affected. Underfoot the leaf litter makes a crackling sound, it is drier and piled up. That means nutrients are not being recycled as efficiently. What effect will that have on the long-term health of the fragment? Will it

gradually run down? The researchers do not know yet, but the project is intended to continue for another 20 years. Along the way, the project should provide many

data to test theories of island biogeography and come up with some practical idea on how groups of Amazon species can be maintained. □

INPE-SPACE INSTITUTE

At the centre of a storm

São Jose dos Campos

It is a curious circumstance that if it were not for a quarrel over deforestation of the Amazon, the National Institute of Space Research (INPE) would be little known abroad. International recognition, or at least the attention of the international media, came to the institute earlier this year when President José Sarney claimed that satellite data analysed at the institute showed that a mere 5.12 per cent of the Amazon had been cut down since 1500. In the quarrel over the validity of the figures that followed (see *Nature* 339, 86; 1989), the world press descended on the institute.

Remote sensing is in fact only one part of what INPE does. It is Brazil's biggest single research institute, employing more than 1,700 researchers in the fields of atmospheric science, astrophysics, space science and meteorology. The institute is responsible for Brazil's satellite programme and will soon see the completion of the nation's first indigenous satellite, SCD-1, being fabricated at the aeronautics company Embraer to INPE's specification.

The 115-kg SCD-1 is a relatively unsophisticated satellite, designed to collect meteorological data from 10 land stations at remote sites across the nation and retransmit them to a receiving station in Mato Grosso. But the satellite is also a key engineering test bed. By 1995, Brazil plans to launch seven satellites for use in telecommunications, search-and-rescue,

meteorology and exploration for mineral resources. That, at least, is the plan.

A major difficulty is that a separate programme to develop rocket launchers being carried out by the Space Activities Institute of the Aeronautics Ministry is way behind schedule. A one-third size scaled-down version of the rocket was launched successfully last May but there is little chance that the full-scale rocket can be ready before mid-1992, more than two years after INPE could have its satellite complete.

The rocket programme has been held up partly because of the unwillingness of the industrialized powers to transfer technology to a programme that is controlled by the military and could easily be diverted for use in long-range ballistic missiles. US refusal to sell inertial guidance systems to Brazil triggered a minor diplomatic row in April. But things are unlikely to change, given that Brazil is an important arms supplier to developing nations and provided rockets, tanks and armoured cars in the recent Middle East conflict.

Many within INPE would like to see a foreign launcher used to put the first satellites into orbit and so keep their part of the programme on schedule. A separate agreement has already been reached to build a remote-sensing satellite for China and launch it on a Chinese rocket. But the military insist that Brazil's programme must await Brazil's own launcher.



Bitter about media treatment, remote sensing director da Cunha.

Meanwhile, Roberto Pereira da Cunha, director of remote sensing at INPE, still feels bitter about press handling of the dispute that thrust prominence on INPE last April. The project was important because it was the first complete satellite image study of deforestation in a decade. But the message was lost amid controversy over the figures.

"The problem is that everyone likes to do sophisticated things", says da Cunha, "everyone likes to talk about the influence of deforestation on climate and global warming but no one wants to do the dirty work of gathering the data. It is a very trivial job for scientists." Deforestation estimates varied from 44 per cent to the World Bank's 12 per cent, but all were projections from old figures. INPE scientists calculated from a mosaic of 101 Landsat images that in fact 5.12 per cent of the government-defined 'Legal Amazon' had been deforested in recent times (not since 1500 as Sarney claimed — there are large areas of very old deforestation).

Arguments continue over whether deforestation should be related to the Legal Amazon, or to the total area of rainforest or some other number but da Cunha says these are not really important.

Absolute values matter little compared with continuing rates of deforestation. He is now trying to obtain funds to study changes through to 1992 as part of the International Space Year, in collaboration with the US National Aeronautics and Space Administration and the European Space Agency.

Da Cunha stresses that remote sensing has many uses and his group, almost 200 strong, puts much of its effort into research in image-processing and generation as well as applications research and technology transfer. Satellite data have been used to settle legal disputes over whether land is in use or not (a key factor in maintaining ownership), to estimate the areas of flooding of the giant reservoirs and to monitor development projects. And in a more esoteric project, some of da Cunha's group are using satellite data to study the gigantic and enigmatic patterns of lines in the desert at Nazca in Peru, to see if they were drawn according to astronomical principles. □



The Amazon basin in mosaic form from INPE. The fine grid of lines at the bottom is the road system in Rondônia. Around it much of the forest has been cut down.

The biggest and the best

São Paulo

It is a cruel shock to the senses to arrive in São Paulo on a wet, dark night. All that can be seen is a spinning confusion of freeways, underpasses, missed turnoffs and high-speed drivers who overtake on the inside and cut in front of one with a little flourish. São Paulo is huge; ringed by factories and slums and dissected by an irrevocably polluted pair of rivers. With more than nine million inhabitants, it is the largest city in Latin America and the third largest in the world.

Surprisingly, dawn reveals that São Paulo is also beautiful. It is a hilly city with plenty of green spaces and views and the broken skyline of tower blocks is never monotonous. Out of town — at around the 20-kilometre mark — the city degenerates sharply into shacks and shabby shanty towns, home to five million of the poorest.

São Paulo is the capital of the richest state in Brazil. Its inhabitants add up to less than a quarter of the total population but they generate half the nation's GNP. This one state contains half of Brazil's top eight universities, and best of them all is the University of São Paulo (USP).

The university is a little out of town, scattered over a slope from the top of which São Paulo is seen spreading across the horizon. At the bottom of the hill is a clock tower and the other usual symbols

rector who shook up the university through insisting on increased emphasis on research.

When he took office, the university had few funds of its own and almost all the money for research, even down to journal subscriptions, came from grants. Now 20 per cent of the university's \$250 million-a-year income from São Paulo state goes on non-personnel costs — laboratories, computers, libraries and the like.

Goldemberg says that when he became rector he decided to "make great efforts to be in the vanguard. Since Brazil needs modern technology we have to put a lot of emphasis on research and high quality." He says he met resistance.

"How do you put pressure on people who all have tenure?" he asks rhetorically. "You use a complex system of incentives — access to money, and better libraries, better communications, computers, sending people abroad . . .". Goldemberg recounts his various temptations and sighs over those who were not ready for them.

"Some areas were not in a position to absorb money. They had no ideas, no contacts. I would tell people to make a list of everything they wanted, then they found they could not make a proper list, they were used to underdevelopment." "People thought I was a bit harsh", he says, "but after the event they thought it was not so bad and even approved of what I had done."

Goldemberg's successor, Roberto Leal Lobo, was chosen by election two weeks ago. As elsewhere, there was heated debate over how much say students and support staff should have in the new 'democratic' election. USP decided on a safe multi-stage election that left the final choice to the professors and, ultimately, the governor.

Lobo is a physicist and was previously vice-rector, but his election was not a foregone conclusion. Running against him were the Secretary of State for Science and Technology (page 371), and half a dozen other university professors. A key issue during the campaign was the university's links with industry. Some candidates were against expansion of industrial involvement, others wanted it to be given the highest priority. Lobo backed a middle way with, he says, "human resources as the first priority, with an overflow to industry. Otherwise we'll just become an industrial research institute."

'Evaluation' was another controversial issue in the election. Many faculty wanted the cosy option of evaluation being left within departments, but Lobo backs the move to more independent assessment. A doctorate is essential now for a staff position and Lobo claims that contracts of

Technology the driving force

Campinas

THE State University of Campinas, located in São Paulo state about 100 km from the capital, is often described as "second only to USP". But it is not an imitator of USP. From the time it began life in 1966, Campinas chose to give priority to science, technological research and, uniquely, links to the production sector. The first courses it offered were in physics, chemistry and mathematics, set alongside technological projects in optical fibres, lasers and computers. Next came engineering and last of all, in the 1980s, the arts.

Campinas now offers a full range of disciplines. It is very popular and has around 23 applicants for every place, making it more competitive even than USP, although it is only half the size. The campus has a quiet feel, set amid sugar cane fields in the rolling countryside a few miles outside the city of Campinas, the second largest in São Paulo state with a population of three-quarters of a million. Although the university was created at the height of the university expansion boom, it managed to keep to a policy of recruiting only the best staff. Sixty per cent of faculty hold doctorates, which is a very high proportion in Brazil, and around ten per cent are from foreign universities.

Campinas faculty like to list their university's unique features. First, almost half of its

non-tenured staff are no longer automatically renewed. Although there is no comparison with the driving pressure of the United States, USP is moving towards a more competitive environment.

Facilities at USP are the best in Brazil and will soon be even better thanks to a \$150 million loan from the InterAmerican Development Bank (IDB). The sophistication of the university allows researchers to keep up with international trends to an extent not visible at other Brazilian universities.

Researchers at the Institute of Physics, for example, were quick to replicate experiments demonstrating high-temperature superconductivity. They also jumped in immediately when they heard about cold fusion, but director Ivan Cunha Nascimento says that "after three months we decided to stick with high-temperature fusion!" The institute has a plasma-physics group and a small tokamak, the only one in Latin America. IDB funds have provided for a new molecular beam epitaxy laboratory and two new electron microscopes.

Elsewhere, the Institute of Oceanography is able to run yearly expeditions to the Antarctic where Brazil maintains an all-year-round base. That gives the geologists from the Institute of Geosciences a

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Goldemberg: centre-stage until January.

of a university campus and, quite out of place, a two-km-long rowing tank. This splendid facility was apparently built for a Pan-American games that was never held.

Ruler of this university for the past four years has been physicist José Goldemberg, a larger than life figure well known both as rector and for his criticism of the government's nuclear programme (see page 374). His term of office ends in January.

The university is big: 35,000 undergraduate students, 15,000 graduate students (half the doctoral degrees in Brazil are taken here), 5,000 faculty and 15,000 support staff. Eighty per cent of students study science, engineering or medicine. Seventy per cent of the faculty hold doctorates, the highest percentage in Brazil.

Goldemberg has been a controversial

14,000 students are pursuing graduate courses. Many are themselves teachers on release from other universities all over Brazil, giving Campinas the reputation of a 'school for schools'. Second, researchers are encouraged to build their own links with industry and even to profit from them. Campinas was the first university to set up a patent office to protect its staff's discoveries (a move which has still found few imitators). Several researchers have set up their own small companies. Third, the university has been aggressive in buying good facilities.

At the end of last year, it bought an IBM 3090, giving it the most powerful super-computer in Latin America. Next, as soon as federal funds are available, will come the National Synchrotron Light Facility. From the beginning it has had the National Center for Research in Telecommunications.

In 1986 it created a large biotechnology institute at the stroke of a pen by buying nearby Monsanto Corporation's Agricultural Research Center. Later, it complemented the centre with a Genetic and Molecular Engineering Center to train people for the expected expansion in biotechnology. Now there is an multidisciplinary programme centred around biotechnology, alongside similar integrated schemes in computers, fine chemicals, energy and new materials. The rector, Paulo Renato Costa Souza, says that his university's philosophy of trying to build integrated programmes, instead of letting everyone submit individual grants, is "completely new in the university system". □

chance to examine structures in Antarctica to which Brazil was once joined. Overall, the impression at USP is of a university that is only occasionally at the international frontiers but is frequently not far behind them.

But like researchers elsewhere, USP staff are held back by petty import restrictions. Some laboratories — the Institute of Chemistry is one — employ several people simply to do the paperwork. But the delay remains. Ana Clara Schenberg, a yeast geneticist at the Institute of Biomedical Sciences, has been waiting for one piece of equipment for four years. It is already obsolete.

Life is particularly difficult for molecular biologists. Even things such as dry ice are in short supply and items that need to be used quickly — short-lived isotopes, for example — are almost unobtainable. And, worst of all, delay destroys flexibility. "If you need to change your research programme, just forget it", says molecular biologist Bianca Zingales. Everything has to be planned ahead around the import of reagents.

Many highly trained young researchers are now returning from abroad, thanks to earlier government decisions to invest in foreign fellowships, offering the chance to push the university ahead. Unfortunately,

Biomedicine with a bite

São Paulo

SNAKES, scorpions and spiders make the Institute Butantan famous. Its serpentarium is São Paulo's most popular tourist attraction and every year thousands of people from all over Brazil send to the institute snakes and other exotic creatures they have found in field or garden.

The snakes are essential for Butantan's function as the main supplier of antivenom in Brazil. Venom is milked from the snakes and used to hyperimmunize horses (the institute keeps 700 at stables outside São Paulo); a specific gamma globulin fraction containing antibodies against the venom is later extracted from the horses' blood. The institute maintains a small hospital, manned 24 hours a day, with a helicopter pad in front of it.

People bitten in the São Paulo region are rushed here — treatment within the first three hours is most effective. Brazil has an impressive array of poisonous snakes and more than 30,000 people are bitten each year.

Snakes are not the institute's only concern. Butantan is essentially a large biomedical research institute joined with a production facility. As well as snake antivenoms, it produces 16 kinds of vaccine. Its 1,300 staff are divided among 20 different departments of basic research and a new Centre of Biotechnology.

Naturally enough, some of its key projects involve improving vaccines. One involves cloning polio virus in vaccinia with the eventual aim of constructing a single-shot vaccine cocktail for polio, measles, hepatitis B, diphtheria, tetanus and whooping cough. Another is to develop a genetically

engineered hepatitis-B vaccine, although the institute may decide to buy the technology from abroad. Projects to produce albumin and factor VIII by new techniques are also under way.

Other groups do basic research on venoms (the important vasodilator drug bradykinin was first isolated here), viruses (it is still relatively easy to discover a new virus a day in the rain forest), the genetics of amphibians and reptiles, and the isolation of natural substances from plants and marine animals.

Techniques developed at the institute are transferred free "to anyone seriously interested in production", says the director, Willy Beçak. "We're not interested in profit but to be in the at the frontier", he says; "the idea is for the institute to be an open place for people from universities; often universities are not the right place for transfer to the production sector."

The institute's popularity rose after it opened a 'reagent supermarket' in an effort to counter the effect of import restrictions.

Special permission to import reagents tax-free in bulk, together with a \$1 million grant from CNPq, enables Butantan to offer immediate delivery of a range of hard-to-import items. But it is not permitted to duplicate reagents manufactured in Brazil. □

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

the returning researchers often seem quite lost — particularly those who have been at top US laboratories.

Around the campus, young people in various stages of 'returnee's syndrome' can be met. Salete Newton, who returned from a postdoctoral fellowship at Stanford University three months ago, is in the fairly early stages. "You don't know what to do", she says, "you know everything will take ten times longer than in America. There the limit is your imagination, here you are trapped by the limitations of the facilities."

A little later, the returnee struggles to see advantages. "At last I have time to read all those papers", says one. "Maybe it's a good thing that I'm forced to think instead of just rushing onto the obvious experiment", says another. Along comes the thought that one is free.

"I realized I could try some stuff that would be thought too risky over there, the sort of thing you would only do if you are already a full professor", says molecular biologist Fernando Reinach, who has no symptoms of the syndrome after three years back home. Returnees also come to realize that at USP one can attract the very best students of one's own.

The fully recovered researcher avoids the most internationally competitive areas but is busy inviting foreign colleagues for short stays (provided they agree to smuggle restriction enzymes, short-lived isotopes and the like through the Brazilian customs), using BITNET and facsimile machines to keep up, building a group in an area of his or her own, and sending off students for doctoral training abroad. Of course, there are always a few who never recover and just settle back. . . . □

Institute earns healthy respect

Rio de Janeiro

THE Oswaldo Cruz Foundation's grounds are ringed with fences, walls and armed guards who harbour deep suspicions about strangers. Presentation of a copy of *Nature* will in no way convince them of one's good intentions. That and the baroque architecture of the 'Moorish Pavilion', set on a hill amid dark trees above, suggests one may have accidentally stumbled on a secret military organization rather than a biomedical foundation.

Later, an explanation emerges. The foundation, on the northern outskirts of Rio de Janeiro, is ringed with *favelas*. A

tute supplies much of the world's yellow fever vaccine.

The production facilities aside, the institute feels little different from a university. But the focus is on "endemic diseases and their vectors", research director Carlos Médicis Morel explains, with the link between research and production clearly in mind. The institute's 200 researchers are expected to publish in international journals and to be capable of true international collaboration — "Not the old kind where we have the diseases and they have the technology", says Morel.

Grants from the World Health Organization provide foreign currency and allow some items to be obtained on a 'diplomatic' basis through the Pan American Health Foundation. The result is a miracle in Brazil — Morel is able to order short-lived isotopes by telex from abroad and see them delivered in 24 hours. "There is a fight for every microlitre", he says.

A new biotechnology department is to combine research and pilot plant production and will concentrate on diseases transferred by the blood — including AIDS, malaria, syphilis and, of course, Chagas' disease.

Chagas' disease, named after the former director, is still both an important research topic at the institute and a major health problem in Latin America. As many as 10–12 million people are affected. Modern medicine has not cured the illness but has brought new modes of transmission. In Brazil, 20,000 people a year are infected through contaminated blood transfusions, perhaps more than the number infected by insect bites. Despite the scale of the problem little was done until AIDS arrived. Previously, poor people were commonly selling blood to private transfusion banks. That is now illegal (thanks in part to the institute's lobbying). But the change has come too late for haemophiliacs. Sadly, 70 per cent of them carry the AIDS virus.

Morel, when not wearing his research director's hat, runs a group developing DNA probes for diagnosis of Chagas' disease. Samuel Goldenberg, a young molecular biologist returned from Paris, is trying a different approach, using recombinant antigens.

Diagnostic kits in use now cross-react with antigens of other diseases. His group has cloned several *Trypanosoma cruzi* surface antigen genes, and one has good potential for use in a test kit.

Goldenberg has a fine sense of the differences between laboratories here and in Paris and the pains and joys of working

in Brazil. His field is the molecular biology of development.

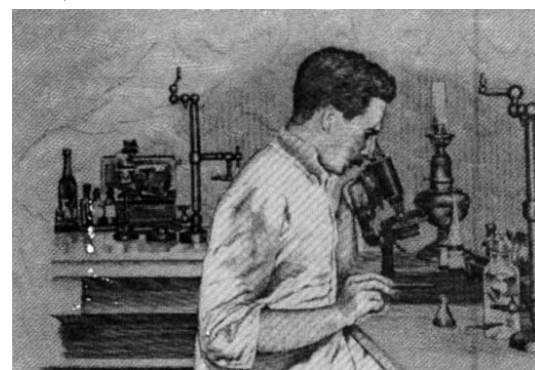
"Elsewhere", he says, "I would probably work with *Dictyostylium*. Here I work with trypanosomes. My main interest was in basic biology, but now we may have an application for our work. When you feel your science is being used for something useful its very very good."

His group has found factors that trigger the switch from the nonreplicative stage of *T. cruzi* to the replicative stage and is looking for enzymes important in the differentiation process. If enzymes unique to the trypanosome can be found, therapy with specific inhibitors may be possible.

Goldenberg's laboratory is overcrowded to bursting point. The corridors are so full of equipment that if two people meet they have to work out a way of backing into the remaining spaces in order to get past one another. If three or more people meet, the moves become just too complex — it is better just to return to one's room.

Goldenberg says that he sometimes "gets really fed up with the difficulties", that "he cannot hire new staff but is simply sending people abroad", that the strict importation rules mean that he is "treated like a smuggler" and that foreign journals come by sea, "three or four months late". But, he says, "the ambition here is not to publish, publish, publish as in the United States — we do want to publish but the main goal is to train more people".

In mid-October, Goldenberg was busy



On the money: Chagas depicted on a 10,000 cruzados banknote.

trying to organize the 16th Annual Meeting on Chagas' disease and leishmaniasis, due to begin a few weeks later at a resort hotel. Six hundred participants were expected. The organizers pay for research students' accommodation, "otherwise the students won't be able to go", says Goldenberg, "and it's a good opportunity for them to meet people, find a place to do a postdoc. But it is a hard task because we cannot find any funds!" When last seen, Goldenberg was listening to the manager of the hotel shouting, "When will you send the money. When will you send the money?" over the telephone. In Brazil there usually seems to be some way to fix the impossible — no doubt the conference went well. □

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Imposing prospect: the Oswaldo Cruz Institute is important politically as well as medically.

few weeks earlier a couple of men dressed as policemen had come along and persuaded a naive academic to loan them his car for an 'exercise'. They then robbed the bank on the foundation's grounds and vanished (or so the story goes).

The normally peaceful institute was set up by Oswaldo Cruz in 1900 to produce vaccines to combat local epidemics of plague and yellow fever. It still produces vaccines (like its sister Institute Butantan, page 369) and over the years has gained a reputation as the most distinguished biomedical research institute in Brazil, with a considerable voice in public affairs.

In its grounds is the National Public Health School which trains most of the nation's health administrators. Akira Homma, the institute's director, is clearly proud that the institute helped to write Brazil's new constitution.

Half the institute's budget comes from the government and a third from the sale of vaccines. Vaccine sales are important, says Homma, because that means "we have our own budget, produced by ourselves which we can invest by ourselves". Production is being built up further with national self-sufficiency the goal. Polio vaccine is still all imported but the insti-

Informatics law and disorder

São Paulo

DÉCIO Leal Zagottis is Secretary of State of the Special Secretariat for Science and Technology of the President of the Republic of Brazil. Had he taken office last year rather than this, he could have done the same job and been simply Minister of Science and Technology; the cumbersome title is the result of a political battle for control of the ministry during the first few months of this year.

The opening broadside came in January with the sudden announcement that the Ministry of Science and Technology was to be abolished and its functions shared out among other ministries.

At the same time, the president agreed to wrenching cutbacks that reduced science and technology budgets by half and would have led to the dismissal of 40 per cent of research staff at CNPq institutes. But the surprise attack was motivated not just by the desire to save money; in the vanguard of the charge were those who opposed the ministry's controversial informatics law.

As it turned out, the abolitionists did not prevail. Congress rallied to the ministry's cause and three months later it was back, albeit under a new title which left it with a little less weight than a full ministry. And Zagottis, whose support for the informatics law is well known, was appointed to head the new agency.

The controversial law was passed in 1984 to foster the local production of computers, software and peripherals by reserving the Brazilian market for 'national' companies, that is, those that are Brazilian-managed and at least 70 per cent Brazilian-owned. The law has been bitterly opposed, both by the United States which sees it as restricting access by its giant computer companies, and by local industry which sees it as forcing up its own costs by making it buy higher-priced lower-quality Brazilian equipment.

Zagottis, a professor of engineering at the University of São Paulo, entertains no doubts about the effectiveness of the informatics policy. Brazil began at the same level as Mexico, but Brazil now has an industry of 60,000 employees, twenty times larger than Mexico's, and supplying over three-quarters of the nation's computers. "Informatics is connected to so many other fields, it is a strategic economic area which must be protected", he says. Protection is intended to come to an end in the early 1990s, to ensure that an inefficient monopoly does not develop.

Zagottis does not think the same policy will work in other high-technology areas. "You cannot have a market reserve policy for biotechnology", he says, "because you cannot do reverse engineering in biotechnology." In other words, while it may be

possible to clone an IBM by 'reverse engineering' (IBM engineers might call it 'copying') it is not possible to clone a genetically engineered drug.

The immediate threat to the informatics law was nullified by the revival of the ministry, but the damage to the science and technology budgets could not be so easily repaired. For almost three months of turmoil, there had been no one to fight for science and technology. Some research institutes, such as the National Institute for Amazonian Research (INPA), that are directly attached to the ministry saw their budget appear in one agency, disappear, appear in another and then disappear again. The result is that they still do not have proper budgets for 1989. Massive cuts for the National Research Council, CNPq, and for the project research agency, FINEP, cannot now be fully restored.

Although it is too early to say what may happen after a new President of Brazil takes office early next year, Zagottis expects that the government will have no long-term investment capacity in the next few years. His solution is to turn to the World Bank and the InterAmerican Development Bank (IDB) to tide Brazil's scientific enterprise over the next next few years. Two big loans are under negotiation: one is a World Bank loan of \$330 million to continue the Programme of Support for Scientific and Technological Development (PADCT), the other an IDB loan of \$1,000 million over five years to support an integrated science and technology plan.

Zagottis says the idea is to use the money to "allow construction of new buildings for research centres, new laboratories and equipment and to train researchers and technicians". The World Bank is, however, not especially pleased with the handling of funds loaned to Brazil for PADCT I, six years ago. An official team sent to examine progress reported at the end of 1988 that if difficulties delaying the projects "are not solved urgently, they may also create serious difficulties in the renewal of the PADCT loans". Sluggish administration, "excessive" and "unjustified" delay in importation procedures, and the "enormous devaluation of allocated resources due to the current inflation rate if the money cannot be spent very quickly" were the principal complaints.

The importation delay for foreign equipment and materials is a terrible

problem. Four-year delays are not uncommon. Zagottis claims that a new law will soon complete changes already under way and ease the import problem. Responsibility for import procedures is being devolved to big universities from CNPq, which was exceedingly slow in handling requests. But there are, as Zagottis says, "problems lower down the hierarchy"; the requests must still pass through CACEX, the import control agency, whose bureaucrats mete out the same treatment on imports worth \$100 or \$1 million and insist on evidence that there is no Brazilian substitute.

But imports are just an irritant compared with the fundamental problem of lack of investment. Brazil spends only 0.7 per cent of GNP — a sum of \$2,500 million — on science and technology, as against the 2.5–3 per cent spent by the industrialized countries (\$150 million in the United States). Zagottis points out that 80 per cent of the money spent by Brazil on science and technology comes from the government, a far higher percentage than elsewhere. If Brazil can boost industry's contribution, then 3 per cent of GNP might be possible for Brazil too. Zagottis backs a short-term rise to 1 per cent of GNP by the government, followed by a steady series of small rises,

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Special secretary for Science and Technology Zagottis supports reserving the Brazilian market for nationally produced computers.

always a little above the rise in GNP. The first goal, he says, is 2 per cent by 2000.

For this to happen, a national plan is needed as well as political stability, not something Brazil seems set to enjoy. Increases in investment by industry will only come from a well thought out programme of tax incentives. At present, it is more profitable for industry to license foreign technology than develop their own. But a failure to act could jeopardize what has been achieved so far. A key strategy has been to send people abroad for advanced training — that is where a large part of CNPq's money goes. As Zagottis says, "There are 7,000 students taking PhD courses outside Brazil; if they don't have a place to come back to they will leave Brazil permanently" and the effort will have been wasted. □

Finance from start to finish

Rio de Janeiro

FINEP, the Agency for Financing Studies and Projects, is an unconventional mix of a conventional research foundation and a conventional development bank. In principle, this gives it a unique power to take a project from the basic research stage, when a government grant is required, through to the development stage when venture capital or low-interest credit becomes important. In practice, things do not work out quite so neatly: FINEP finances a heterogeneous brew of projects, all at different stages, ranging from university research on AIDS to an industrial company's attempt to develop the first fully Brazilian automobile.

FINEP is one of the few federal agencies to have avoided the move to Brasília, something which clearly pleases its staff. Its offices are in downtown Rio de Janeiro.

Support for basic research comes principally through FINEP's role as executor of the National Science and Technology Development Fund (FNDCT), worth \$107 million in 1987. The fund provides for infrastructure, new buildings and new equipment. Significant extra help comes through \$180 million of World Bank funds, most of it matched by Brazilian funds, in the Programme of Support for Scientific and Technological Development (PADCT). Grants are given to support projects, to upgrade and build new facilities and to support university graduate courses. Support for individual re-

searchers is left to the National Research Council, CNPq.

"Earthquakes and turmoil" is how 1989 was described at FINEP. Thanks to the savage cuts of the Summer plan and the temporary disappearance of the Ministry of Science and Technology, to which FINEP is answerable, the agency will be lucky to receive 60 per cent of its 1987 budget. The World Bank has also been frank in its criticism of the slow progress in giving out PADCT funds. Of the \$180 million loaned to Brazil in 1983, less than half had been disbursed by 1989. Import regulations and excess bureaucracy — little of it of FINEP's making — have been the major problems.

FINEP claims a good track record. When it was set up in 1967 to develop Brazil's engineering industry, there were very few companies, foreign or Brazilian. According to João Luiz Coutinho de Faria, president of FINEP, "twenty-two years later we have 200 companies; most are Brazilian with an annual turnover of \$1,450 million".

FINEP has 700 staff, more than 100 with doctoral or masters' degrees, and last year supported 1,840 projects split almost equally between scientific research and technological development.

Coutinho de Faria points out that demand for the technological component was more than \$1,000 million last year. He hopes to see FINEP's funds return to their 1987 level of \$300 million next year and climb quickly to the \$800 million mark. □

CAPES

Supporting graduates by algorithm

Brasília

POSTGRADUATE programmes throughout Brazil are coordinated by a Ministry of Education agency called CAPES; not surprisingly, few know the agency by its lengthy title of "Coordinating Agency for Advanced Training of High-Level Personnel". CAPES gives grants and fellowships to 14,000 students at Brazilian universities and to 2,000 students in North America and Europe.

Overall, CAPES provides backing for 700 graduate courses; two-thirds of all the courses in Brazil. It also runs a programme to help build up institutions so that they can offer new courses.

The graduate-training system grew rapidly in the 1970s during the expansion of the university system. At that time, many young and inadequately-trained staff obtained jobs. Today, the boom is over but there are still many unproductive faculty and graduate courses are of very uneven quality. CAPES makes sure it helps the good courses but is sometimes criticized for not being tough enough on the bad.

CAPES' director, José Ubirajara Alves, says that graduate courses are now the "best-evaluated sector in the science and technology system". He is a rare type in Brasília, describing himself as just a "technician" who has come up with an "algorithm" to ensure that money goes to the courses that deserve it. In fact, Alves is a very well-respected mathematics professor with plenty of experience of starting graduate courses from nothing. When he joined the university at Fortaleza in his home state of Ceará in 1965 he was the only person on the staff with a doctorate (from the University of California, Berkeley). He established its first graduate course — the first at a university north of Brasília. Fortaleza now has twenty courses.

Alves' algorithm takes into account the standard of the course, the numbers and types of professors and students, and the area (some types of research are more expensive than others). When the calculation is complete the budget goes to the university to provide for a quota of students, selected by the university. □

Fast forward for grants

São Paulo

WHEN scientists complain that the National Research Council (CNPq) is too slow, or too bureaucratic, or too opaque, or too arbitrary, they often end by saying, "Why can't it be like FAPESP?" FAPESP is São Paulo state's own research foundation, receiving by law 0.5 per cent of state income each year to spend on scientific research.

FAPESP is fast: it normally makes decisions on grants within ten weeks of application, less than half the time required by CNPq (or by US NSF).

To ensure openness, applicants are given as much information about the progress of their grants as needed, a job which keeps a couple of secretaries busy full time. And once a grant is given, FAPESP will help to compensate for changes in exchange rate that push up the cost of imported materials. The foundation runs on 50 staff, against CNPq's 1,000.

Alberto Carvalho da Silva, FAPESP's director, is modest about the foundation's reputation, pointing out that CNPq has a budget twelve times as big as FAPESP's \$42 million, and has to deal with unexpected cuts and political uncertainties. But he is critical of CNPq's staff numbers, saying that there are so many employees that it has to create something for them to do. Carvalho da Silva has been a victim of the workings of bureaucracy himself; he was purged by the military from his university post in 1969.

FAPESP's operations are much the same as those of research councils everywhere. Support goes to scholarships (not more than 50 per cent of the total), research awards, a small number of large long-term projects (most recently, support for participation in BITNET), visiting professorships, publications, meetings and so on. One novel programme provides 500 scholarships a year to encourage students to move into research by providing support for a pre-graduation research project.

Expansion and change are on the way. The new state constitution increases FAPESP's budget to one per cent of state tax income and gives it new responsibility for the support of technological research. Until now, 90 per cent of the budget has gone to basic research at the universities. In the future there will be a new commitment to the private sector on training and pre-competitive research. "Thus far, we've been run by scientists", says Carvalho da Silva, "now we'll be looking for international advice. We will use our contacts with the United States and the United Kingdom to look at science parks and adapt their models to our realities". □

The council has the last word

Brasília

It is a mean trick to arrive at the headquarters of the National Research Council (CNPq) after several weeks of accumulating complaints from Brazilian scientists about their parent research organization. CNPq is too cumbersome, they say, it is supposed to handle import procedures but everything takes years, travel grants arrive after conferences are over, foreign collaborations founder in paperwork, and grants are spread too thin, or go to foreign scholarships so that nothing is left for those at home.

And yet more, CNPQ swallows you up complains one University of São Paulo researcher, "if you criticize them, they say 'come and help us', and you end up being involved in all those difficulties too".

Guilherme Brandão, Special Assessor for International Cooperation at CNPq, is a pipe-smoking meditative type who only occasionally shows signs of mild irritation at some of the more unreasonable behaviour of scientists. CNPq is not, in any case, a complete master of its own destiny, but has to weather all kinds of political storms and bureaucratic meddling.

Science lives in a difficult environment. "Science and technology are not thought of as a high priority", says Brandão, "we are pressured by so many other things we need to solve, we have immediate problems in health and nutrition and so on". And because there is no tradition of research and development by industry (expenditure on research by industry is just 5 per cent of the total), science is totally dependent on the government. "When there is a government financial crisis, we suffer badly", says Brandão. One of those crises arrived earlier this year when budgets were cut by half and, for a short while, it seemed that some 40 per cent of the staff of CNPq's research institutes might be without a job.

To summarize, CNPq is a research council just like others elsewhere, except that along with its grants programme it has eleven research institutes (see page 357) and a large scholarship programme. It has been very powerful, with full responsibility for science and technology policy. But that responsibility went to the Ministry of Science and Technology (now the Special Secretariat for Science and Technology) in 1985. CNPq has never really relinquished its claim to the area and despite lacking an official mandate, continues work at its Centre for Policy Studies.

If scientists complain of tardiness, they are perhaps insensitive to CNPq's scale. Each year, 40,000–50,000 applications for grants come in to CNPq's offices (including travel grants and so on), along with over 100,000 applications for scholarships. Suitable reviewers have to be found

for each application, CNPq can provide for a third of the applicants in each category.

CNPq employs 1,000 staff but even so Brandão says that sometimes "the amount of paperwork drives us crazy here". The system is to be rationalised next year. Instead of four fixed decision periods a year, a system which can never generate a decision in less than 4–5 months, CNPq will deal on a once-a-year basis with some things (scholarships, for example) and continuously with others that need immediate decisions (such as travel grants). That change should do away with a substantial fraction of the complaints of CNPq's inflexibility.

CNPq's scholarship programme is big, too big says critics, who think more of the resources should be directed to ongoing research rather than training. Ten years ago there were just 5,000 research scholarships in total. Today there are 30,000 inside Brazil and 3,500 overseas, part of a definite policy that top priority should go to human resources. But the number abroad will lessen next year as people return and efforts are made to save foreign currency. Inside Brazil the number will increase to provide 44,500 places. A scholarship is not riches — a doctoral student receives around \$600 a month — but thanks to "a very great joint victory of CNPq and CAPES", the amounts are tied to staff salaries so providing compensation for inflation.

On the extraordinary difficulty of
CAPITAL CITY

Back to the future

BRASILIA, the futuristic city of the 1950s, is laid out in the shape of a gigantic bow with its central arrow an enormous, thousand-yard wide avenue that, at its tip, juts out into an artificial lake. Along it, at great intervals, are the television tower, various monuments, the cathedral, the double row of identical, rectangular blocks containing the ministries, and then Congress and the Supreme Court.

The bow is a curving set of parallel roads along which are organized the city's 'sectors': the hotel sector, the banking sector, the cultural sector, the hospital sector, the embassy sector, the sector of individual housing . . . and so on. The great unifying architectural plan continues to inspire in areas of astonishing simplicity of design, where traffic flows swiftly past rows of simple, white concrete apartment blocks raised on stilts to allow clear sight lines. Elsewhere, the road plan remains but the city has filled in a medley of styles. CNPq has its offices in the sector of "autarquias" — autonomous organiza-

importing equipment and reagents for research, Brandão wearily says, "Yes, we understand the community's complaints. But sometimes they complain in the wrong place". CNPq is caught up in a much vaster bureaucracy.

All imports must pass through CACEX, and CACEX is essentially an instrument of foreign policy. It is necessary to show that no similar Brazilian product exists but 'similarity' is often a "complicated business", says Brandão. Next a pro-forma invoice is necessary, then "when you send this to CACEX they may have a problem with foreign currency and freeze everything".

While the difficulties are undoubtedly real, some problems may remain at CNPq. According to the World Bank, FAPESP, the São Paulo-state research council (see opposite) takes from "six months to a maximum of one year to import equipment", while CNPq takes "one and a half to two years". And CNPq took "at least 45 days to log in importation requests" (it has since improved).

CNPq has to spend a lot of time fighting off arbitrary decisions from elsewhere. Brandão relates how in January the government suddenly decided that money sent abroad for training grants should be exchanged at the tourist rate rather than the official rate, immediately cutting the grants' value by more than half. In the ensuing havoc, many students decided not to take up their scholarships. The decision was reversed only a few weeks ago.

One complaint which Brandão clearly feels unfair is that Brazil is trying to shut out foreign researchers or limit their access to the Amazon. Brandão says that 99 per cent of foreign expeditions (of which there are 30–40 a year) are approved, but there have to be formalities to ensure that an expedition will not intrude on Indian or private land and so on. Collaborative programmes, of which there are thousands, are organized through Brazilian counterparts. Brandão says the system ensures that Brazil obtains some benefit from research conducted here.

CNPq is looking to the future and looking for innovative ways to boost research investment by industry. It is not easy because in many industrial sectors a foreign licensing agreement captures a sure product while home development involves risk. But in other, newer sectors, foreign companies do not want to license their technology. CNPq is trying to get Congress to adopt flexible tax incentive strategies. In the meantime it is introducing a new programme of human resources in strategic areas (biotechnology, new materials and the like), to support the training of personnel in any institution, private or public, provided it is tied to a long-term research and development programme. The goal is to have 3,500 people in the programme by next year.

Some projects from the past . . . and future

São Paulo

WHEN the first wave of oil price surges arrived in 1973, Brazil reacted with two new and massively expensive programmes to reduce dependence on energy imports. One was intended to provide electricity from nuclear power, the other to substitute sugar alcohol for petroleum. Both projects were creations of a military government which could exercise strong central control over resources. Fifteen years later, both projects are floundering as they face markets which demand a more rational allocation of resources.

Opposition to the nuclear programme was strong from the start. "It was very clear that nuclear energy for electricity production was a complete misconception," says University of São Paulo (USP) rector, José Goldemberg, a well-known critic of Brazil's nuclear plans, "seven billion dollars gone with little to show".

A small US reactor is in operation but the agreement signed with West Germany to build eight reactors has yet to create one operating plant. The programme, "was slowly delayed but never cancelled," says Goldemberg, "scientists opposed it because it involved very little technology transfer. Industrialists who lost opportunities because of support for the programme opposed it." Last week, West Germany extended the agreement with much reduced goals. But the situation may change again under a new president. Goldemberg believes that "in a democratic society it will become increasingly difficult to justify the expenditure against social programmes". The sugar alcohol programme appears very much more successful.

Four million of Brazil's 12.5 million vehicles run on alcohol produced from fermentation of sugar and massive reduc-

tions in petroleum imports have been achieved. But the programme is in serious economic and political trouble. Brazil has found large oil reserves of its own. Sugar alcohol is proving more expensive to produce than petroleum, the distilleries are producing huge quantities of polluted water, and subsidies for sugar-cane growing distort local economies by encouraging farmers to grow cane even when vegetables are in short supply.

These, at least are the arguments advanced by the state petroleum industry, which wants to see petrol roll back some of alcohol's advance. With the backing of the Ministry of Mines and Energy, the petroleum industry has scored well in the media war. The sugar alcohol industry, backed by the Ministry of Industry and Commerce, is in retreat but it too claims that subsidies disguise the true state of the petroleum industry. The battle of the subsidies will be fought between the two ministries under the new president and in the minds of consumers, who must decide which kind of car to buy.

One factor in the battle is the scope for improvement in sugar alcohol production. Waste treatment is already possible and biotechnology may offer much more. Little sustained effort has gone into the improvement of yeasts. Genetic engineering can permit new substrates to be used, such as starch which yeast cannot normally degrade. In a clever piece of engineering, Ana Clara Schenberg at USP has already constructed yeast strains carrying α -amylase genes from mouse, plus other yeast genes introduced by conventional genetics and shown the new strain can produce ethanol directly from starch. While the project is at the laboratory stage, it suggests that it is too soon to begin writing off sugar alcohol. □

Parallel secrets

FEARS that Brazil is intending to develop nuclear weapons in secret are fading with the return of democracy.

In 1987, after the disappearance of the military regime, a secret 'parallel' nuclear programme emerged with the announcement by President José Sarney that the navy had succeeded in low-level enrichment of uranium fuel. The work was carried out at the Institute for Energy and Nuclear Research (IPEN), adjacent to the USP campus.

The aim of the programme is not a bomb, but fuel for a nuclear-powered submarine. Although Brazil has no enemies and the expense of developing nuclear submarines seems superfluous, the official reply is that the project will take 20 years and by then Brazil will be a great power. The navy was much impressed by the performance of nuclear submarines during the Falklands conflict.

As the 'parallel' programme went public, demands increased for greater accountability of the military's activities. The new constitution prohibits the development of nuclear weapons and gives Congress control over the nuclear energy programme. But the constitution left possible the development of 'peaceful' nuclear explosives. The first steps towards civilian control came last year when the government established a Superior Council for Nuclear Energy as an advisory board to the president. Goldemberg was one of three civilian scientists on the council, which includes 17 ministers. Although the meetings are not frequent, Goldemberg says that "the president attends and it is now very difficult to do anything under cover".

Next step could be a congressional committee with oversight of military programmes. Brazil would then be close to the certainty that the submarine programme could not be diverted towards manufacture of a bomb. □

Rio de Janeiro

IF new styles of collaboration between the universities and industry are what is needed to boost investment in science, then the Federal University of Rio de Janeiro (UFRJ)'s island may be the place to look for future trends. Here Antonio Paes de Carvalho has confidently launched a new biotechnology venture, Bio-Rio, which links industry, the Oswaldo Cruz Foundation and the university. Within ten years, Paes de Carvalho confidently predicts there will be 70 companies here, employing some 4,500 people and enjoying \$250 million in sales.

Paes de Carvalho is a cardiac electrophysiologist, an unusual skill from which to move into biotechnology. But in 1980, he says, he was director of the Institute of Biophysics at UFRJ and became "worried about the financing of research". As there were experts in tissue culture and molecular biology in the institute, Paes de Carvalho thought biotechnology might be a good bet.

So with the belief that "you have to do things yourself", he set up a non-governmental, biotechnology foundation close by. The foundation will operate central research facilities and allocate surrounding land, at a nominal cost, to private enterprise. The central facility will offer specialized services (for example, peptide and nucleotide sequencing and synthesis) and provide 'incubator' space where projects can be developed.

In October the central laboratory facility was being carved out of the vast kitchen (conveniently complete with cold storage rooms) of an abandoned university cafeteria. This month, the first two companies are moving into their incubator space.

There are other biotechnology initiatives elsewhere but none based on direct interaction between industry and university. "The biggest fight," says Paes de Carvalho, "is to convince the university that this kind of industrial proximity is OK". The university authorities at first rejected the project, but were eventually persuaded with help from the rector — as he was a communist he apparently could not be accused of selling out to capitalists.

The government has invested \$3 million in Bio-Rio, and industry another \$1 million. Paes de Carvalho expects industry to provide another \$100 million over five years, encouraged by big incentives from the National Development Bank. Critics say the plan's weakness is the shortage of qualified researchers, most of them in universities and without interest in links to industry.

But Paes de Carvalho believes that he can engineer the interaction between university and industry so that it will be painless for researchers and not threaten basic research. □

Upper bounds on 'cold fusion' in electrolytic cells

D. E. Williams*, D. J. S. Findlay†, D. H. Craston*, M. R. Sené†, M. Bailey†, S. Croft†, B. W. Hooton‡, C. P. Jones*, A. R. J. Kucernak*, J. A. Mason§ & R. I. Taylor*

* Materials Development Division, † Nuclear Physics and Instrumentation Division and ‡ Nuclear Materials Control Office, Harwell Laboratory, UK Atomic Energy Authority, Didcot, Oxfordshire, OX11 0RA, UK

§ Reactor Centre and Centre for Fusion Studies, Imperial College, Silwood Park, Ascot SL5 7PY, UK

Experiments using three different calorimeter designs and high-efficiency neutron and γ -ray detection on a wide range of materials fail to sustain the recent claims of cold fusion made by Fleischmann *et al.*¹ and Jones *et al.*². Spurious effects which, undetected, could have led to claims of cold fusion, include noise from neutron counters, cosmic-ray background variations, calibration errors in simple calorimeters and variable electrolytic enrichment of tritium.

(ref. 1 and M. Fleischmann, personal communication) (Fig. 1). We found these to be inaccurate instruments with some very subtle sources of error which it is necessary to appreciate and analyse in detail. We used sixteen such cells, containing different-size cathodes (1-, 2-, 4- or 6-mm Pd rods) and different electrolytes (0.1 M LiOD, 0.1 M LiOH, 0.1 M NaOD or 0.1 M NaOH). Figure 2a,b shows the results obtained for a typical cell. An immediately evident characteristic is the sloping baseline, with the sawtooth pattern a consequence of the regular refilling of the cell. The baseline slope can be quantitatively accounted for by a variation of the calibration constant, k , with the level of liquid in the cell. This is a result of radiative losses through the vacuum jacket⁹ and conduction up the glass inner wall of the cell¹⁰ (see Fig. 2 legend). A consequence of the sloping baseline is that any calibration is only valid for a particular liquid level. We regularly refilled these cells to a reference level at which the calibration had been performed. We calibrated the cells with the heaters before the electrolysis was started and again much later in the run, during the electrolysis, as is evident in Fig. 2a (see Fig. 1 legend for details of the calibration procedure). There was no statistically significant difference between the different calibration sets, so all data were combined.

Figure 2c shows results for two of the calorimeters at the calibrated liquid level, in the form of percentage excess of apparent output power over the Joule input power. There was apparently an endothermic period at the beginning of the run, whose duration increased with cathode diameter. We speculate that this was due in part to poor stirring of the solution during hydrogen uptake by the cathode which caused a temperature gradient in the cell (gas is not at first evolved at the cathode and the larger diameter cathodes were shorter, to keep a constant surface area). An analogous effect at the start of the electrolysis was reported by Lewis *et al.*⁸, who showed that the apparent heating coefficient of a similar cell varied during the first part of a run. Because of this effect, we excluded the first 10,000 minutes of data from each cell when calculating the statistics. Table 1a shows the mean absolute power deviation for each cell and the standard deviation in this value. The standard deviation for all of the H₂O cells was not significantly different from that of the D₂O cells (F test, 1% level) so that the data from all of the cells can be used to estimate the error: $\sigma = 0.048$ W (that is ± 5 –10%). A more detailed analysis of this error¹⁰ showed it to be largely determined by the variability in the liquid level after refilling the cell, but also with a contribution from unquantified variations in the heat loss by conduction up the calorimeter wall to the air above the water bath. Compared with these errors, effects that resulted from temperature gradients inside the cells were minor. The design used here varied from that in ref. 1 in that the glass sleeve that supported the calorimeter and which was in direct contact with the inner wall, provided a large area of thermal contact with the water bath, and thus reduced the effect of the ambient temperature variations.

As we expected, occasional points from both H₂O and D₂O cells lay outside the 'control limits' of $\pm 2\sigma$ (Fig. 2c). No points lay outside $\pm 3\sigma$. No cell showed two or more consecutive points outside $\pm 2\sigma$, and the number of points lying above the control limit in the D₂O-cell experiments was no different from that in

RECENT publications^{1,2} reporting electrochemically induced nuclear fusion, at room temperature, have aroused great interest. The signatures reported are excess heat output, neutron emission and tritium generation from cells with palladium cathodes¹, and neutron emission alone at a much lower level from cells with titanium cathodes². Conventional nuclear physics predicts that fusion between light nuclei requires either very high temperature (as in a tokamak) or unusually close proximity of the two nuclei (as in muon-catalysed fusion). The calculated fusion rate at the internuclear separation in the deuterium molecule (0.74 Å) is $\sim 3 \times 10^{-64} \text{ s}^{-1}$ (ref. 3), so the rates reported in ref. 2 ($\sim 10^{-23} \text{ d-d-pair}^{-1} \text{ s}^{-1}$; d is deuteron) are not easy to understand in the context of the known interstitial-site separations (~ 3 Å), even more so when it is realized that this rate is a severe underestimate because of incomplete deuterium loading in titanium (see Materials characterization section). The reaction rates reported in ref. 1 are even more difficult to understand⁴, and the existence, despite strong arguments to the contrary (ref. 5, for example), of an unknown mechanism, which results both in an extraordinary enhancement of the reaction rate and a suppression of the normal nuclear-reaction channels, has been postulated.

In the first reports of the electrolytic cold fusion effect it was stated that the effect is not consistently reproducible, and that it both takes some time to appear and that it may subsequently disappear. Because electrochemical phenomena can be sensitively affected by the state of the surface, some irreproducibility is not, in itself, surprising, and other recent reports (refs 6–8, for example) give well documented accounts of failures as well as successes (R. A. Huggins and A. J. Appleby, Workshop on Cold Fusion, Santa Fe, May 1989) in observing the claimed effects. The clear lack of reproducibility necessitates significant replication, with controls at least equal in number to the number of tests, if positive results are to be viewed with confidence, and exploration of many different, well characterized, material and electrolyte combinations. Particularly, the timescales and achievable concentrations for electrolytic loading of deuterium into palladium and titanium, the quantities of hydrogen 'impurity', and the species detectable by surface analysis of used cathodes ought to be determined.

Calorimetry

We used three types of calorimeter. First, we built calorimeters of size and design similar to those used by Fleischmann *et al.*

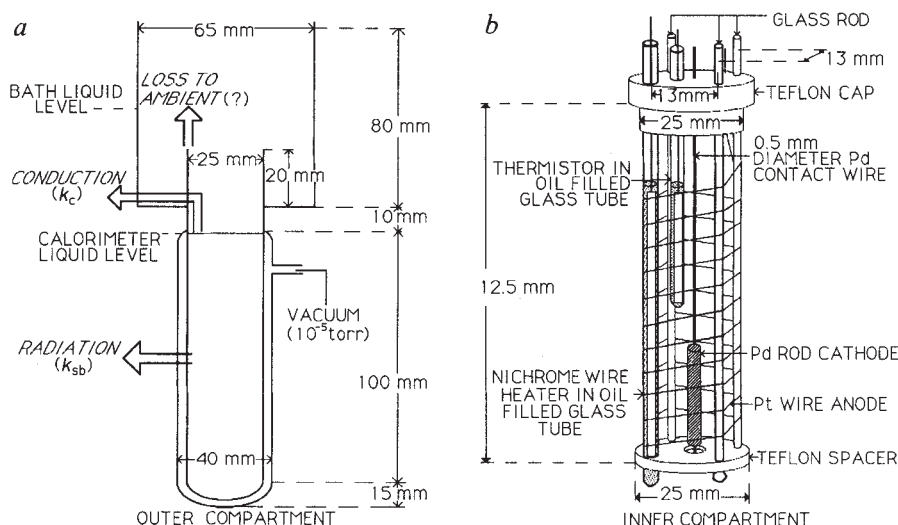
the H₂O-cell experiments. Therefore, we conclude that, within the experimental error, there was no significant anomaly in the behaviour of the D₂O cells compared with that of the H₂O cells. If, however, the mean power deviation (Table 1a) for all eight of the D₂O cells ($+9.91 \pm 0.47\%$ or 0.021 ± 0.004 W; mean Joule input 0.96 W) is tested against that for all eight of the H₂O cells ($-2.46 \pm 0.72\%$ or -0.018 ± 0.004 W, mean Joule input 0.81 W for LiOH cells and 0.61 W for NaOH cells), it is clearly highly significant. Because no sequence of individual points lay outside the control limits established above, we suspect that there is another, unknown source of error that scales with the input power. This, of course, could be the postulated 'fusion effect', but the magnitude of the effect is commensurate with the errors. The only reliable way of checking this is to construct calorimeters

that are free of major sources of errors, and in particular, do not have a sloping baseline. We therefore also used both isothermal and steady state (heat flow) calorimeters which satisfy this criterion.

We have also analysed in detail the slope of the output curve¹⁰ to look for momentary power pulses on timescales shorter than the interval between calibrated points. If the Joule input remains constant then this slope should not vary. Any sudden or momentary extra power input, q_c , would change the slope by approximately $(k_{sb} + k_{c,0})q_c/M_0$ (M_0 is the water equivalent of the calorimeter, other symbols defined in Fig. 2 legend). The number of significant deviations from the mean slope was found to be roughly the same for both the H₂O and the D₂O cells. For the whole data set, the largest power excursion for a D₂O cell (6-mm

FIG. 1 Schematic diagram of the FPH-type heat-flow calorimeter used here. Heat flow paths are indicated here and discussed further in Fig. 2 legend.

METHODS. As well as the Fleischmann, Pons and Hawkins (FPH) type, we used¹⁰ two other calorimeter types—an improved heat-flow calorimeter (IHF) and an isothermal calorimeter. The IHF calorimeters differed from the FPH type in three important ways. First, they were larger: a 500-ml-capacity cylindrical vessel constituted the cell, with anode-cathode spacing ~ 2 cm. Second, the electrolysis vessel was inserted, using a film of oil for thermal contact, into a tightly fitting aluminium can which was itself packed around with insulating material and placed in a Dewar flask. Whereas the temperature of the cell contents in the FPH calorimeter was determined using a glass-clad thermistor immersed in the cell itself, in the IHF design the temperature of the aluminium can was measured: the can defined an isothermal surface for conduction of heat away from the cell and its use as the temperature measurement surface eliminated the sloping-baseline problem of the FPH design. Third, the space within the Dewar flask above the electrolysis vessel was filled with a polystyrene cap extending well above the Dewar flask, with the aim of significantly reducing unquantified heat losses to the atmosphere. The two types of heat-flow calorimeter were operated in water baths held at a constant temperature of 20°C ($\pm 0.08^\circ\text{C}$). The tops of the water baths were covered with a polystyrene lid. In both cases the cells comprised a spirally wound Pt wire anode (0.25 mm) and a central Pd cathode (Johnson Matthey). For the FPH cells we prepared the different electrolytes using either conductivity water (H₂O cells) or slightly tritiated D₂O (specific activity 13 kBq ml^{-1}) of isotopic purity initially $>99.9\%$. We used 0.1M LiOD (D₂O from Aldrich, measured $>99.9\%$ initial isotopic purity) as the electrolyte in the IHF calorimeters. The Pt and Pd contact wire was shrouded in glass tubing in the IHF cells to prevent any possible catalytic recombination of the electrolysis products. In later experiments with the FPH cells we also used screened electrode contacts, but this had no effect on the results obtained. The IHF cells contained a Pd wire electrode insulated to the very tip, which was positioned about three-quarters of the way up the cell. This was used to define the internal liquid level during filling, or refilling, of the cell. The larger volume of electrolyte in the IHF cells (300 ml) meant that refilling was required only occasionally, but the large thermal mass meant that the response was slow (time constant ~ 12 h): thus they were not sensitive to small bursts of heat. We calibrated the heat-flow calorimeters using a nichrome wire heater placed in an oil-filled glass tube which was in contact with the cell contents. The calibration procedure involved operating the heater, either without electrolytic current, or with a current significantly less (0.2–0.4 times) than the normal electrolysis current, until a steady-state temperature was attained. With the FPH cells, use of a lower electrolysis current ensured continued stirring while diminishing the errors arising from the baseline drift; before each individual calibration the cells were refilled. An empirical calibration curve was thus obtained by fitting the observed thermistor resistance, R , to a range of applied powers, P , using the equation $P = a - b(\log R) + c(\log R)^2$, where a , b and c are the fitted parameters. The standard error of estimate on the fitted curve varied from 10 to as much as 70 mW with the FPH cells, largely reflecting the error in extrapolation to



the reference level. Neutron counting was performed in conjunction with the operation of the FPH cells, using two independent banks of counters mounted above the cells, on top of the water-bath cover¹⁰. Detection sensitivity above background was 3 event s^{-1} in the cell. We never observed a signal above background. The isothermal calorimeter (J. A. Mason, R. W. Wilde, J. C. Vickery, B. W. Hooton and G. M. Wells, Proc. 29th A. Meet. Inst. Nuclear Materials Management, Las Vegas, Nevada, June 1988) comprises three concentric aluminium cylinders each separated by a heat transfer medium with a relatively low thermal conductivity. The cylinder temperatures are maintained by electrical heaters wound as helical coils around each cylinder. Temperature control is achieved by resistance thermometers on each cylinder which are used in conjunction with classical control software and the cylinder heaters. The rate of thermal energy evolution in the measurement chamber (12.5-cm diameter, 26 cm high) is determined by measuring precisely the electrical power supplied to the chamber. We operated the calorimeter at a measurement-chamber temperature of $42 \pm 0.001^\circ\text{C}$. The measurement-chamber power resolution is $<5\text{ mW}$ for an operating power of 20 W. The electrolytic cell contained ~ 1 l of 0.1M LiOD. The cathode was contained in a perforated glass canopy to prevent the evolved gases from mixing inside the measurement chamber. The anode was a Pt foil cylinder 3 cm high and 12 cm in diameter surrounding the cathode. The cell was thermally coupled to the calorimeter measurement chamber by conducting oil. Measurements using Pt cathodes and also using nichrome wire heaters showed that there was a small systematic error in the calorimeter, an apparent power excess that varied linearly with the input power up to 100 mW for an input of 15 W. A linear fit to these measurements (12 points, standard error of estimate 8.8 mW) was therefore used to apply a correction to the apparent excess measured for the Pd cathodes. The measured output power of both the isothermal and IHF cells was corrected for the power loss that is due to evaporation of the electrolyte, assuming that the electrolysis gases were saturated in water vapour as they passed out of the cell (25.0 mW A^{-1} at 42°C): $q_v = (p_v/p_0)(1.5)\Delta H_v/(2F)$ where q_v is the power loss, p_v denotes the saturation vapour pressure and ΔH_v the latent heat of evaporation of the electrolyte, p_0 is the atmospheric pressure, F is the Faraday constant and I is the current (assumes ideal gases).

TABLE 1 Calorimetry results

(a) FPH Calorimeters: current 300 mA, current density 80–110 mA cm⁻², 0.1 M LiOH, NaOH, LiOD, NaOD

Pd ^a			H ₂ O Cells		D ₂ O Cells ^b	
Diameter	Volume	Cation	Mean Joule input ^c	Excess ^{d,e}	Mean Joule input ^c	Excess ^{d,e}
Surface area	Charging time		(mW)	(mW)	(mW)	(mW)
6 mm	0.28 cm ³	Li	820	-45 ± 50	960	12 ± 57
3.9 cm ²	870 h	Na	590	14 ± 71	980	57 ± 44
4 mm	0.19 cm ³	Li	800	-19 ± 28	960	-20 ± 29
3.5 cm ²	810 h	Na	610	-55 ± 53	960	-11 ± 49
2 mm	0.094 cm ³	Li	780	40 ± 53	930	61 ± 26
3.1 cm ²	740 h	Na	620	-12 ± 38	940	36 ± 40
1 mm	0.071 cm ³	Li	830	-46 ± 64	980	43 ± 64
2.8 cm ²	670 h	Na	620	-18 ± 60	980	-4 ± 35
Mean excess	$\pm(\sigma/\sqrt{n})$ (mW) ^f		-18 ± 4 (n=143)		+21 ± 4 (n=146)	

(b) IHF and isothermal calorimeters: 0.1M LiOD

		Pd						
Calorimeter type	Type Surface area	Total charging time Volume	Current (mA cm ⁻²)	Time (h)	Joule input ^c (W)	Excess ^d (mW)		
IHF	2-mm rod ^a 3.1 cm ²	797 h ⁱ	156	231	2.341	78 ± 77 ^m		
		0.16 cm ³	156	206	2.380	27 ± 48		
			219	187	3.842	-1 ± 58		
			279	173	5.188	-55 ± 101		
	Cast ^g 14 cm ²	797 h ⁱ	35	236	2.172	100 ± 40		
		0.88 cm ³	36	201	2.219	47 ± 36		
			49	187	3.489	-31 ± 51		
			62	173	4.863	-120 ± 86		
	Isothermal	2-mm rod ^a 1.3 cm ²	284 h ^o					
			0.063 cm ³	159	284	0.925 ^o	-10 ± 12 ⁿ	
Cast beads ^h 5 cm ²		355 h ^o	50	7	0.985	33 ± 12		
		0.5 cm ³	30	15	0.415	11 ± 11		
			70	7.8	1.796	5 ± 12		
			20	15	0.214	10 ± 11		
			100	8.6	3.405	-5 ± 12		
			120	270	4.825	2 ± 13		
			40	8.1	0.675	7 ± 12		
			80	5.7	2.311	9 ± 12		
			200	4.8	12.314	-10 ± 12		
			160	4.1	8.140	-17 ± 12		
			220	4.8	14.675	-9 ± 12		
		Melt-spun- ribbon ⁱ 74 cm ²	74 h ⁱ	14	74	8.991	64 ± 17	
			0.35 cm ³					
			2-mm rod ^j 1.3 cm ²	323 h ⁱ	152	70	1.091	7 ± 11
			0.066 cm ³	530	253	10.358	15 ± 14	
		8-mm bar ^k 28 cm ²	520 h ⁱ	30	520	8.974	36 ± 15	
1.5 cm ³								

^a Johnson Matthey (JM) 'Specpure'; drawn from sintered stock prepared from high-purity powder.^b D₂O from Harwell reference stock, contains 13 kBq ml⁻¹ tritium.^c Mean Joule input power supplied to cell (see Fig. 2 legend). Values for IHF and isothermal calorimeters have been corrected for heat loss that is due to evaporation (see Fig. 1 legend).^d Excess power = measured cell output power - calculated Joule input power.^e (Mean ± 1σ) of values calculated after each refilling to the reference level, excluding first 10,000 minutes of polarization (see text).^f Mean and standard deviation of the mean calculated for all H₂O data points and all D₂O data points.^g Specially produced material supplied by JM—prepared from cast Pd stock that was argon-arc melted into rod form using a gravity casting process. The rods were subsequently sliced and bent to decrease the loading time required (maximum distance from bulk to surface ~1 mm). The sample was cleaned using acetone, 10% HCl and distilled water.^h 'Specpure' Pd arc melted three times under argon on a water-cooled copper hearth.ⁱ A variety of ribbons prepared (JM) by melt spinning of cast or sintered Pd. A proportion of the ribbons were heat treated (JM) for 20 min at 100 °C under 10% H₂/N₂. Ribbon thickness, 125 μm.^j Type ^a that was subsequently vacuum degassed at 1,200 °C, and loaded with deuterium at a pressure of 40 bar. The sample was cooled to liquid-nitrogen temperature before transferring to the calorimeter to minimize loss of D₂.^k Sintered high-purity bar, sliced and bent to decrease the loading time required (maximum distance from surface to bulk ~1 mm).^l D₂O from Aldrich Chemical Co., contains ~15 kBq ml⁻¹ tritium.^m Mean and standard deviation of all data points after temperature stabilization: data point every 3 min.ⁿ Standard deviation given by $\sigma = \sqrt{\sigma_b^2 + \sigma_y^2 + \sigma_c^2}$ where σ_b is the standard deviation of the baseline measurement, σ_y that of the power measurement with the cell running and σ_c is the standard error of estimate of the correction line (see Fig. 1 legend). σ_b and σ_y were typically 6 mW, σ_c was 8.8 mW.^o Small error in this particular measurement gave rise to the apparent small endotherm.

rod) was ~ 45 mW and for an H_2O cell (6-mm rod) ~ 40 mW. These are small compared with the input power. It is certainly clear that no unusually large power pulses occurred. Given the difficulties in operating these calorimeters, very occasional occurrences of small fluctuations cannot be considered as support for a 'fusion' hypothesis.

The other two types of calorimeters used in this study are described in Fig. 1 legend and results are given in Table 1b. We explored a range of different preparations of palladium and of current density, up to nearly 600 mA cm^{-2} . With the improved heat-flow calorimeters (IHF), the sign and magnitude of the power excess varied with Joule input power, but was always $< 5\%$. Expressed in terms of the volume of the palladium cathode, this sets a limit of $100\text{--}500 \text{ mW cm}^{-3}$. The most accurate calorimeter used was the isothermal calorimeter (minimum-detectable power change ~ 10 mW; minimum-detectable energy in any brief burst ~ 40 J). We analysed these results at four-hour intervals by averaging both the Joule input power and the measured output power over a period of ~ 20 min. Inspection of the data collected every four minutes between the regions of analysis showed no obvious signs of any short heat 'bursts' and we found no trend with time of the measured output power under any of the conditions used. Table 1b shows that we obtained thermal balance to better than 20 mW ($24\text{--}240 \text{ mW cm}^{-3} \text{ Pd}$). We observed slight thermal excesses ($30\text{--}60 \text{ mW}$) during the initial charging period of the palladium beads, and during runs with high-surface-area cathodes at high current: it can reasonably be assumed that a small amount of recombination (4% at most) was responsible for this effect.

Neutron counting

We investigated the emission of neutrons from a wide range of cells using three different detector systems (Table 2). The large, high-efficiency detector with which most of the neutron measurements were made is an oil-moderated assembly of $56 \text{ }^{10}\text{BF}_3$ proportional counters configured as 5 concentric rings^{11,12}. The total efficiency for 2.45-MeV d-d neutrons is 44%. We built an automatic cell shuttle mechanism to exchange regularly two nominally identical cells, only one of which was powered. This enabled the background (which is due mostly to cosmic rays, there being no anti-coincidence counter arrangements) and any signal to be counted virtually simultaneously. In operation the cells were exchanged every 5 min, and the data from the 5 rings were recorded separately. Data from a typical run are shown in Fig. 3 as differences in the count rate between the powered and unpowered cell. Although in the particular example shown, two spikes can be seen in the count rate differences from the detector as a whole, it is clear that these spikes are due entirely to the misbehaviour of ring 4, and are therefore spurious.

The details of the cells and results of the measurements are given in Table 2. The lowest limits (2σ) on neutron emission derivable from Table 2 for palladium are $1.5 \times 10^{-2} \text{ n s}^{-1} \text{ g}^{-1}$ or $1.5 \times 10^{-2} \text{ n s}^{-1} \text{ cm}^{-2}$, if the emission were sustained over the whole run. For titanium, the values are $3 \times 10^{-3} \text{ n s}^{-1} \text{ g}^{-1}$ and $4 \times 10^{-3} \text{ n s}^{-1} \text{ cm}^{-2}$, although the latter limit is too high because it does not apply to the runs using granules having a large and indeterminate surface area. If we assume the emission to be sustained over only a one-hour period at most, then the limits are, for palladium: $7 \times 10^{-2} \text{ n s}^{-1} \text{ g}^{-1}$ or $4 \times 10^{-2} \text{ n s}^{-1} \text{ cm}^{-2}$, and for titanium: $8 \times 10^{-3} \text{ n s}^{-1} \text{ g}^{-1}$ or $2 \times 10^{-2} \text{ n s}^{-1} \text{ cm}^{-2}$ (much less in the runs with granules). The neutron-emission rate limits are several orders of magnitude below the rates of $\sim 10^4 \text{ s}^{-1}$ reported in ref. 1 and about one order of magnitude below the rate of $\sim 0.4 \text{ s}^{-1}$ reported in ref. 2. It is significant that the limits in our work were also obtained for cells in which cold fusion could be expected to be enhanced, in particular by using titanium cathodes in the form of a large (40 g) mass of porous granules to increase both the reaction volume and the surface area, and an electrolyte specifically chosen to promote deuterium loading into the cathode¹³ (see Fig. 3). As discussed later, we have not

expressed the results in terms of the number of deuterium pairs within metal lattices.

Because the net count rates given in Table 2 are differences between much larger background rates, runs were undertaken with a calibrated ^{252}Cf source ($0.024 \pm 0.003 \text{ fissions s}^{-1}$) emitting $0.09 \pm 0.01 \text{ n s}^{-1}$ and a physically identical blank. The result was $0.14 \pm 0.05 \text{ n s}^{-1}$, in satisfactory agreement. Furthermore, the expected neutron output of the unpowered UPT_3 cell, which is due to spontaneous fission of ^{238}U , was 0.22 s^{-1} . The value measured in comparison with an empty cell was $0.20 \pm 0.06 \text{ s}^{-1}$. Our confidence in the results of the shuttle differences seems justified. Measurements on CeAl_2 and UPT_3 cells were included in the hope that high effective electron masses corresponding to these metal crystal lattices would mimic in some way the fusion enhancement effect of 'heavy electrons' such as muons in binding the deuterium nuclei closer together. The palladium-ribbon run with the cell power switched on and off every two hours (which was about eight times the characteristic diffusion time in the ribbon) was undertaken to enhance the appearance of non-equilibrium effects, and results (negative) are given in Fig. 4.

γ -ray counting

As it has been shown that cold proton-deuteron (p-d) fusion is expected to proceed at rates greater by ~ 8.5 orders of magnitude than d-d fusion³, we carried out an alternative investigation of electrolytic enhancement of hydrogen-isotope fusion by looking for the $\text{D}(p, \gamma)^3\text{He}$ 5,488-keV γ -rays from p-d fusion. We used a lead-shielded 113-cm^3 n-type high-purity germanium (HPGe) crystal γ -ray spectrometer to search for any γ -rays in the energy range 0.1–7 MeV emitted by a variety of cold-fusion cells operating with a mixture of light and heavy water. The results, also given in Table 2, are consistent with no γ -ray emission. The lowest limits (2σ) on γ -emission derivable from Table 2 are, for palladium, $7 \times 10^{-3} \gamma \text{ s}^{-1} \text{ g}^{-1}$ or $1 \times 10^{-3} \gamma \text{ s}^{-1} \text{ cm}^{-2}$, and for titanium $1.4 \times 10^{-3} \gamma \text{ s}^{-1} \text{ g}^{-1}$ or $4 \times 10^{-4} \gamma \text{ s}^{-1} \text{ cm}^{-2}$, if we assume that the emission is sustained over the whole run. Jones *et al.*² report that the fusion activity may last for only 4–8 h and begin ~ 1 h after the cell is powered, and in these circumstances typical values for the standard deviation in the γ -ray emission rates are $\sim 0.05 \text{ s}^{-1}$ and the 2σ limits for titanium are $4 \times 10^{-3} \gamma \text{ s}^{-1} \text{ g}^{-1}$ or $3 \times 10^{-3} \gamma \text{ s}^{-1} \text{ cm}^{-2}$. Post-run analysis of the hydrogen isotopes taken up by the palladium cathodes with the mixed light and heavy-water electrolytes used gave very satisfactory D:H ratios, lying in the range 1.5–2.5. We note that if the proposed enhancement of the fusion process is as valid for p-d as for d-d fusion then, because of the enhanced tunnelling in the lighter d-p system³, the γ -ray measurements actually provide a much more stringent limit on the cold fusion process than do the neutron measurements. The interpretation of the results from some of these cells was complicated by the dissolution of gold anodes and the consequent gold deposition onto the cathode: this also would have affected the original work².

Tritium enrichment

Fleischmann *et al.*¹ claimed a tritium production rate of $\sim 10^4 \text{ atom s}^{-1}$, commensurate with their reported neutron emission rate. They used a differential technique in which the tritium accumulation in a cell with a palladium cathode was compared with that in a cell with a platinum cathode. They took samples for analysis at regular intervals and maintained a constant total electrolyte volume by the addition of fresh D_2O .

Electrolytic enrichment is widely used to increase the concentration of tritium in water before analysis¹⁴. Reproducible results require careful control of the electrolysis, as the enrichment factor can vary widely (refs 15, 16 and R. L. Otlet, personal communication): important effects are seen with change of electrode materials, with variation in the condition (activity) of the electrode surface, with the current density (overvoltage) and

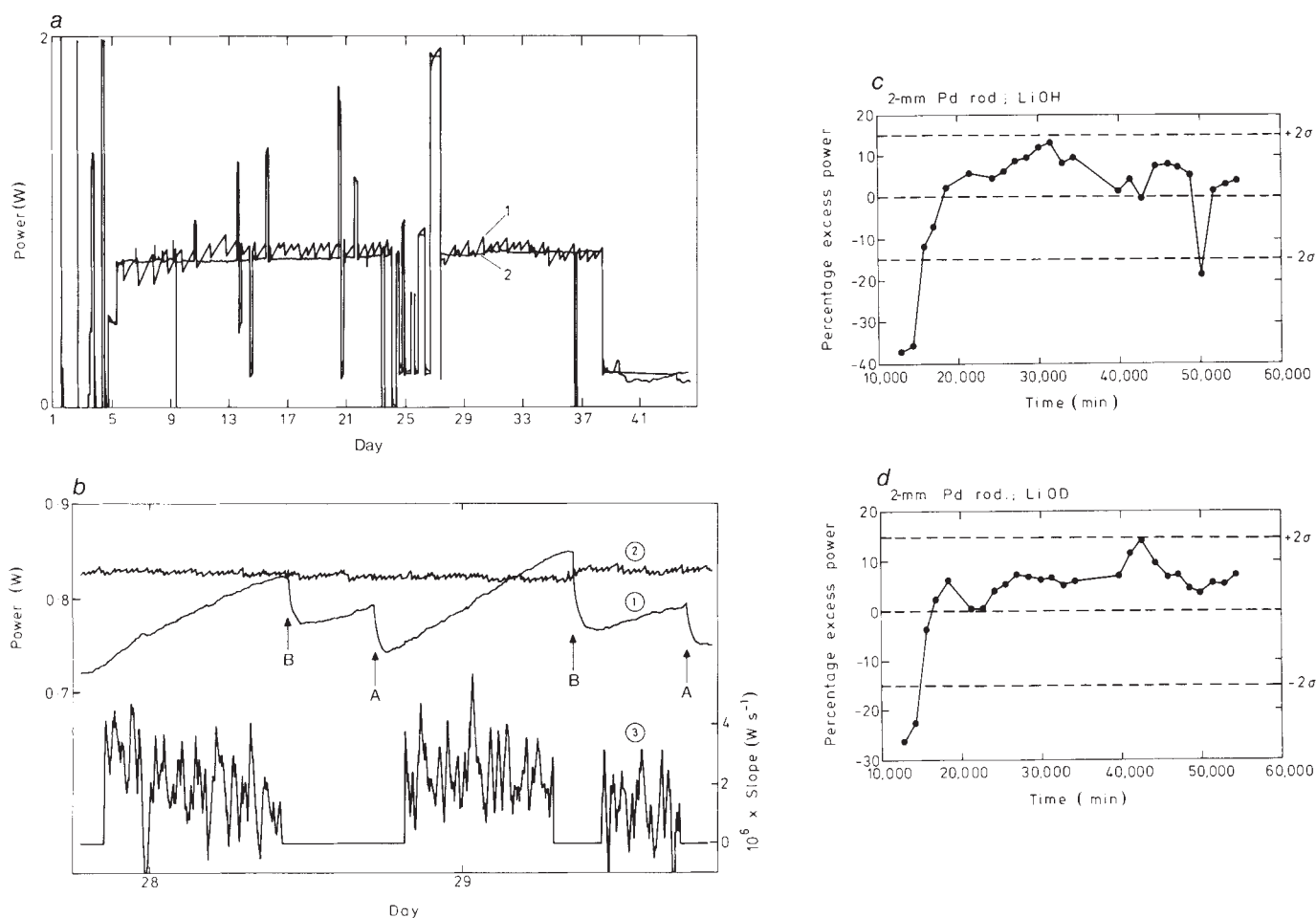


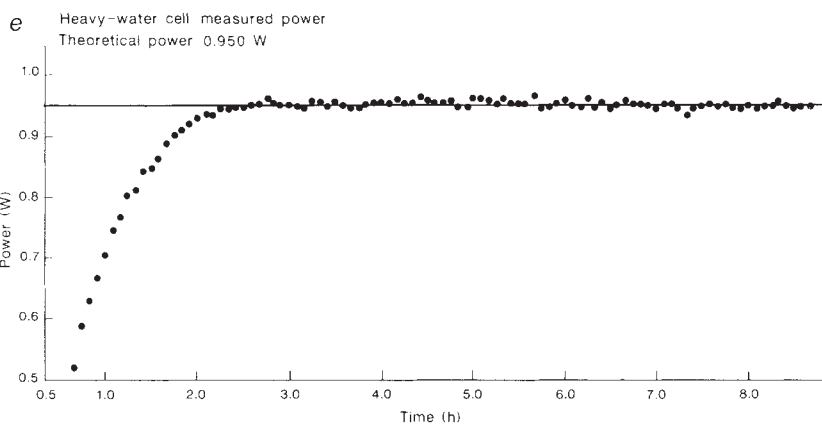
FIG. 2 *a*, Raw data from FPH-type calorimeter containing a 4-mm Pd rod (1.5 cm long, Johnson Matthey 'specpure', drawn from sintered stock) in LiOH electrolyte. Line 1 represents the output power calculated from the thermistor reading and line 2 represents the Joule input power to the cell, $P_{in} = I(V - V_0)$ where $V_0 = \Delta H_d / 2F$ (1.527 V for D_2O and 1.481 V for H_2O , ΔH_d being the enthalpy of dissociation, for example, $D_2O(l) \rightarrow D_2(g) + \frac{1}{2}O_2(g)$). The large step variations are calibrations. *b*, An expanded region of *a*, which emphasizes the sloping baseline. Lines 1 and 2 are as in *a*. Line 3 is the gradient of the apparent output power calculated by differentiation of the data using a seven-point Savitzky-Golay routine²³, stepping one point at a time. At points A the calorimeter was topped up to the reference mark with H_2O pre-warmed to the cell temperature. At points B a volume of liquid estimated from the electrolysis rate was added. *c* and *d*, Results for the two most consistently exothermic FPH calorimeters—0.2 × 3-cm Pd rod in (1) LiOD and (2) LiOH—in the form of percentage excess of apparent output power over the Joule input power. The error bars represent the control limits $\pm 2\sigma$ calculated from the results obtained for all the different cells (see text). *e*, Raw data obtained using the isothermal calorimeter (20 × 2-mm diameter Pd cathode), immediately following the application of the input power, with the line representing the Joule input power and the dots the observed cell power. The response time of this calorimeter is governed by the time to obtain thermal mixing of the cell contents, and is faster than that of the simple calorimeters, despite the large solution volume.

SLOPING BASELINE. The response of the 'simple' FPH calorimeter can be modelled by¹⁰

$$P_{in} = k\Delta T = (k_{sb} + k_c)\Delta T \quad (3)$$

where⁹,

$$k_{sb} = \sigma A_1 [T_1^4 - T_0^4] \approx 4\sigma A_1 T_0^3 [1 + 3\Delta T / (2T_0)] = k_{sb0} [1 + 3\Delta T / (2T_0)] \quad (4)$$



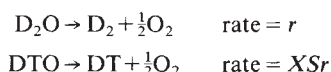
and¹⁰

$$k_c \approx \kappa A_2 / l = \kappa A_2 / (l_0 + V_m \delta t / (2F\pi r_i^2)) \approx k_{c0} (1 - \alpha \delta t)$$

(P_{in} is the Joule input power, k is the calorimeter constant, k_{sb} and k_c are the contributions to the calorimeter constant of radiative losses through the vacuum jacket and conduction up the glass inner wall respectively, σ is the Stefan-Boltzmann constant, A_1 is the contact area of the solution with the wall of the cell, T_0 is the bath temperature, T_1 is the cell temperature, $\Delta T = T_1 - T_0$, κ is the thermal conductivity of the glass, l is the distance from the liquid to the point where the inner glass wall comes into contact with the bath and its initial value is l_0 , A_2 is the cross-sectional area of the inner glass wall, I is the cell current (300 mA), V_m is the molar volume of cell solution, δt is the elapsed time after refilling the cell, r_i is the internal radius of the glass vessel and $\alpha = V_m / (2F\pi r_i^2 l_0)$). We measured the calorimeter constant for the cells to be $\sim 0.1 \text{ W K}^{-1}$, with the calculation using equation (4) indicating roughly equal contributions from k_{sb} and k_c . A more complete description of the calorimeter should also take into account the effects of the solution loss on the heat capacity of the calorimeter. It can be shown however that the contribution of this to the sloping baseline is insignificant¹⁰.

with the temperature. Without precautions, the variations in enrichment factor can be more than a factor of two¹⁶. The claims therefore need to be assessed against this known variable electrolytic enrichment. Surprisingly, Fleischmann *et al.*¹ claimed that there was no electrolytic enrichment in their platinum cell.

If the rates of electrolytic evolution of hydrogen isotopes are written as follows



where r denotes the electrolysis rate, S the tritium-deuterium separation factor and X the mole fraction of tritium, T, in the solution, then the tritium accumulation is given by¹⁰

$$\frac{X}{X_0} = \frac{1}{(S + \alpha + \gamma)} \left\{ 1 + \alpha + \gamma - (1 - S) \exp \left[-\frac{(S + \alpha + \gamma)rt}{N_0} \right] \right\} \quad (1)$$

where the total solution volume is maintained constant by the addition of fresh D_2O (containing a mole fraction X_0 of tritium, directly proportional to the disintegration rate), the ratio of the sampling rate of the solution to the electrolysis rate is α and γ denotes the ratio of the rate of evaporation of the solution to the electrolysis rate—a small correction that may be calculated assuming that the electrolysis gases passing out of the cell are saturated with water vapour. On the timescales of interest, variability in S would give a variable enrichment

$$\delta \left(\frac{X}{X_0} \right) = -\frac{rt}{N_0} \delta S \quad (2)$$

If the solution becomes contaminated by hydrogen absorption from the atmosphere, then if S_{obs} denotes the apparent DT enrichment factor derived from the application of equation (1), $S_{\text{obs}} = S/[0.5(1+f)]$, where $f = (1 - X_{\text{H}})/(1 + X_{\text{H}}(S_{\text{HD}} - 1))$ with X_{H} denoting the mole fraction of hydrogen and S_{HD} the HD separation factor (relative rate of reaction of HDO and D_2O). Therefore, as well as the inherent variability from one electrode to another, it is evident that any variability in the amount of hydrogen pickup will give a variation in the apparent enrichment factor:

$$\frac{\delta S_{\text{obs}}}{\delta X_{\text{H}}} = \frac{S_{\text{HD}}}{2[1 + X_{\text{H}}(S_{\text{HD}} - 1)^2]}$$

We found that, unless exceptional precautions were taken (and we believe that we used experimental procedures very similar to those used in ref. 1 (M. Fleischmann, personal communication)), values of $X_{\text{H}} \approx 0.07$ were common. Under these conditions, a change of X_{H} of only 0.01 would give a change in S_{obs} of 0.02, which could be significant given the smallness of the claimed effect.

Any assessment of whether differential enrichment can be considered to account for the results in ref. 1 depends critically on the value of X_0 , which was not reported. Using values¹ of r (1.24×10^{18} atom s^{-1}) and N_0 (14.6×10^{23} atom) we calculate from equation (2) that if X_0 were at the extreme low end of the range for commercial D_2O (3 Bq ml^{-1}), a value of $\delta S = 0.46$ would be required; if it were moderate ($10\text{--}15 \text{ Bq ml}^{-1}$) a value $\delta S \approx 0.1$ would be needed; if it were high (80 Bq ml^{-1}) $\delta S \approx 0.02$ would suffice.

The applicability of equation (1) was confirmed experimentally, on both platinum and palladium cathodes, in conjunction with calorimetric and neutron-counting experiments (see Fig. 5 legend) using D_2O with an initial tritium content of 13 kBq ml^{-1} (efficiency corrected). The fit of this equation to all the experimental data (Fig. 5) gave, for palladium, $S_{\text{obs}} = 0.59$, and for platinum, $S_{\text{obs}} = 0.61$. Correction for the uptake of hydrogen, using $S_{\text{HD}} = 6$ (ref. 16), gave $S \approx 0.48$, in agreement with previous work and theoretical expectations ($S = 0.46 \pm 0.02$, (refs 14, 15 and D. S. Rawson and R. L. Otlet, personal communica-

tion)). For individual electrodes, S_{obs} for palladium varied from 0.46 to 0.65 (corrected S from 0.42 to 0.58), whereas S_{obs} for platinum varied from 0.56 to 0.85 (corrected S from 0.43 to 0.69)—a total range for individual electrodes of $\delta S_{\text{obs}} = 0.4$, which is enough to account for the results of Fleischmann *et al.*¹

It is clear from these results and discussion that more evidence needs to be presented before the tritium accumulation reported¹ can be considered as experimentally reliable evidence for the occurrence of a fusion process.

Materials characterization

With $125\text{-}\mu\text{m}$ palladium foils, hydrogen loadings (determined¹⁰ by hot-extraction mass spectrometry of specimens frozen in liquid nitrogen, and independently by electrochemical extraction) of $\text{H/Pd} = 0.95 \pm 0.05$ were achieved in $\sim 1 \text{ h}$ at 100 mA cm^{-2} in 0.1 M LiOH solution with a platinum anode. The limit of deuterium loading in 0.1 M LiOD was lower (0.84 ± 0.03). Rods (1- and 2-mm diameter) polarized for extended periods in neutron counting and calorimetry experiments showed $\text{D/Pd} = 0.76 \pm 0.06$. It is well known¹⁷ that the equilibrium pressure for a given deuterium loading is higher than that for the same hydrogen loading. Current-interruption methods confirmed that overpotentials in the range $0.8\text{--}1 \text{ V}$ (ref. 1) were being obtained. From an initial composition $>99.9\%$ D_2O , the solutions degraded to $\sim 98.5\%$ in 24 h and to $88\text{--}98\%$ (analysis by infrared spectrometry) following electrolysis for many weeks. The resulting ratio H/D in the palladium was $0.02\text{--}0.04$.

During electrolysis, all of the palladium cathodes became covered with a layer that varied in appearance from a dull tarnish to a dense jet black. In the latter case, loose black material was also formed, which in extreme cases came off during electrolysis, resulting in quite heavy erosion of the cathode. The layer itself evidently represented a modification of the morphology of the cathode at the surface. The formation of a thick black layer was enhanced at high current density, at high temperature, on smaller diameter wire and by frequent abrupt alterations of the current density repeated over a long period. The layer was more noticeable on the outside of a spiral-wound cathode than on the inside. The layer was also more noticeable on cathodes polarized in D_2O than on those polarized in H_2O (perhaps this is related to the greater equilibrium gas pressure for equivalent composition in the Pd-D system), and was different in appearance on materials from different sources. These observations may be explained by the old idea¹⁷ that microfissures, or rifts, develop in palladium to release the mechanical strains resulting from the heavy loading of hydrogen, together with the assumption that any such effect would depend on the stress state of the metal surface and its microstructure.

Lithium was present in the surface layer on cathodes used in LiOD , and analysis by secondary-ion mass spectrometry (SIMS) apparently showed a concentration profile extending about $1 \mu\text{m}$ into the metal. SIMS images showed that the lithium was not uniformly distributed, however, and it seemed likely that it was trapped in microfissures in the surface layer.

Surface analysis showed a number of other species on and in the surface layer, notably small quantities of platinum and traces of copper, zinc, iron, lead and silicon: platinum would have originated from the anode and silicon from the glass container. No doubt the majority of the other contamination would have come from the solution: the levels found (a few atom per cent, confined to the surface layers) were consistent with deposition by extended electrolysis from a solution of concentration around $10^{-9}\text{--}10^{-10} \text{ M}$. One possible criticism is that low levels of such deposition could poison any essential catalytic activity: we therefore used pre-electrolysis in several experiments, in an attempt to lower the surface contamination. We either made repeat experiments on the same solution, simply changing the cathode, or treated the solution beforehand in a separate cell.

Because of claims that an unusual mechanism might lead to

TABLE 2 Measured neutron and γ -ray emission rates for cold fusion electrolytic cells

(a) Neutron emission rates

Material	Cathode		Electrolyte	Anode	Details of electrolysis			Measured neutron yield ^a (n sec ⁻¹)
	Mass (g)	Surface area (cm ²)			Current (mA)	Typical voltage	Duration (h)	
Pd rod 2 mm ^a	3.4	5.7	0.1M LiOD	Pt wire	360	5	914	These cells monitored initially by low-efficiency n-detectors with lower detection limit of ~ 100 s ⁻¹ . Later moved to a dual-cavity neutron detector ²⁰ with lower detection limit ~ 2 s ⁻¹ . No detected n-emission.
Pd wire 1 mm ^a	0.94	3.1	0.1M LiOD	Pt wire	200 ^f	4	916	
Pd wire 1 mm ^b	0.94	3.1	0.1M LiOD	Pt wire	200	4	916	
Pd wire 1 mm ^b	1.4	4.7	0.1M LiOD	Pt wire	300 ^g	5	917	
Pd wire 1 mm ^c	0.47	1.6	0.1M LiOD	Pt wire	750	11	856	
Pd plate 1 mm ^d	7.5	13.5	1M LiOD	Pt sheet	2,000	10	307	
Pd foil	0.5	8.0	0.1M LiOD	Vitreous C	400–120 ⁱ	4–20	142	
Pd/22Ag ^e	1.6	3.8	0.1M LiOD	Pt wire	380	5	859	
Pd/22Ag ^e	1.6	3.8	0.1M LiOD	Pt wire	380 ^u	5	547	
Pd foil ^f	0.075	1.0	0.1M LiOD	Pt wire	30	3	87	
Pd foil ^g	0.075	1.0	0.1M LiOD	Pt wire	250	7	6	-0.019 $\pm$ 0.042
Pd foil ^h	0.075	1.0	0.1M LiOD	Pt wire	250	7	16	0.00 $\pm$ 0.08
			+17 mM Na ₂ S					0.04 $\pm$ 0.11
Pd foil ⁱ	0.075	1.0	0.1M LiOD	Pt wire	250	8	16	0.005 $\pm$ 0.063
Pd foil ^j	0.075	1.0	0.1M LiOD	Pt wire	250	8	16	0.051 $\pm$ 0.061
Pd foil ^k	0.075	1.0	0.1M LiOD	Pt wire	1,000	15	56	0.068 $\pm$ 0.061
Pd foil	0.63	8.0	0.1M LiOD	Pt foil	1,000	14	17	-0.110 $\pm$ 0.082
Pd foil	0.63	8.0	0.1M LiOD	Pd foil	1,000	18	22	0.012 $\pm$ 0.079
Pd ribbon ^l	0.45	0.92	0.1M LiOD	Pd foil	600 ^v	15	110	-0.063 $\pm$ 0.096
Pd foil	0.60	8.0	0.1M LiOD	Au wire	1,000	12	43	0.068 $\pm$ 0.058
Pd pellet ^m	3.0	2	0.1M LiOD	Pt foil	650	18	68	0.007 $\pm$ 0.042
Pd pieces ⁿ	4.4		0.1M LiOD	Pd foil	500	8	88	0.012 $\pm$ 0.033
Pd pieces	4.4		0.1M LiOD	Pd foil	500	15	88	0.003 $\pm$ 0.046 ^x
Ti foil	0.038	1.0	0.1M D ₂ SO ₄	Au wire	250	5	34	-0.091 $\pm$ 0.054
Ti foil	0.038	1.0	0.1M D ₂ SO ₄	Au foil	1,000	5	3	-0.22 $\pm$ 0.12 ^y
(continued)			0.1M D ₂ SO ₄		1,000	5	4	-0.22 $\pm$ 0.12 ^y
			+0.02M PdCl ₂					
Ti rod	2.2	4.3	Jones ^q	Au foil	100	4	3	-0.06 $\pm$ 0.15 ^y
(continued)					500	6	2	-0.42 $\pm$ 0.21 ^y
Ti granules ^o	0.5		0.1M D ₂ SO ₄	Pt wire	250	7	22	0.00 $\pm$ 0.072
Ti granules	0.5		0.1M D ₂ SO ₄					
			+3 mM Na ₄ P ₂ O ₇	Pt wire	250	7	66	0.044 $\pm$ 0.054
Ti granules	40		0.1M D ₂ SO ₄	Pt wire	250	25	22	-0.009 $\pm$ 0.058
Ti granules	40		0.1M D ₂ SO ₄					
			+3 mM Na ₄ P ₂ O ₇	Pt foil	250	25	22	-0.010 $\pm$ 0.060
Ti granules	17		Jones	Au foil	500	26	24	0.021 $\pm$ 0.054
Ti granules								
Ti/6Al/4V	2.8	15.0	0.1M D ₂ SO ₄	Pt wire	670	5	87	-0.003 $\pm$ 0.031 ^x
Ti/6Al/4V ^p	2.7	14.3	0.1M D ₂ SO ₄	Pt wire	660	5	19	0.031 $\pm$ 0.044 ^x
TiFe granules	1		0.1M D ₂ SO ₄	Pt foil	250	8	24	0.12 $\pm$ 0.06
CeAl ₂ granules	11		0.1M LiOD	Pt foil	500	22	22	-0.051 $\pm$ 0.056
UPt ₃ granules	57		0.1M LiOD	Pt foil	500	20	94	0.10 $\pm$ 0.06

(b) γ -ray emission rates

Material	Cathode		Electrolyte ^z	Anode	Details of electrolysis			Measured 5,488-keV γ -ray yield ^{aa} (γ s ⁻¹)
	Mass (g)	Surface area (cm ²)			Current (mA)	Typical voltage	Duration (h)	
Pd foil	0.58	7.5	0.1M LiOH	Pt foil	450	9	45	(-4.9 $\pm$ 7.43) $\times 10^{-3}$
Pd foil	0.74	9.5	0.1M LiOH	Pt foil	510	9	164	-0.011 $\pm$ 0.004
Pd sheet	7.7	12.9	0.1M LiOH*	Pt foil	200	4.6	50	-0.017 $\pm$ 0.012
Pd rod	3.1	2.3	0.1M LiOH	Au foil	1,000	26	17	0.008 $\pm$ 0.011
Ti foil	0.089	2.0	Jones*	Au foil	320	5	42	-0.022 $\pm$ 0.011
Ti foil	2.2	32	Jones	Au foil	1,000	4	12	-0.015 $\pm$ 0.034
Ti foil	2.2	32	Jones	Au foil	1,000	4	6	0.012 $\pm$ 0.007
Ti granules	25		Jones	Au foil	710	13	8	0.032 $\pm$ 0.036
Ti granules	25		0.1M H ₂ SO ₄	Au foil	400	20	25	-0.019 $\pm$ 0.018
Ti granules	25		0.1M H ₂ SO ₄	Au foil	765	16	16	0.017 $\pm$ 0.022
UPt ₃ granules	42		0.1M LiOH	Pt foil	300	31	166	(3.1 $\pm$ 6.0) $\times 10^{-3}$

^a Cells provided by M. Fleischmann. Cathodes analysed for H, D after use—results H/D: 0.01, 0.02 D/Pd: 0.84, 0.72.^b Cathode wound into a tight spiral. Examined initially in the large n-detector, without shuttle, for 147 h with estimated detection limit 1 n s⁻¹ and a further 5 h with detection limit 0.2 n s⁻¹. Surface of cathode then rubbed with S before further use.^c Cathode cleaned with emery paper after 300 h.^d Cathode examined initially in the large n-detector, without shuttle, in 0.1M LiOD for 183 h; estimated detection limit 1 n s⁻¹. Abraded with 400-grit SiC paper before use.^e Tubes from D storage system for Tokamak.^f Cathode analysed for H, D after use—results: H/Pd: 0.03 D/Pd: 0.80.^g Electrolyte of the above run reused.^h Cathode and electrolyte of the above run reused; Na₂S added as concentrated aqueous solution.ⁱ Foil vacuum degassed 1,000 °C before use. Analysed for H, D after use—results: H/Pd: 0.02 D/Pd: 0.83.^j Foil dipped in Na₂S (concentrated solution) before use. Analysed for H, D after use—results: H/Pd: 0.05 D/Pd: 0.80.^k Cathode cut in half and analysed for H, D after use—results: H/Pd: 0.02, 0.02 D/Pd: 0.85, 0.78.^l Melt-spun ribbon provided by Johnson-Matthey Technology Centre.^m Arc remelted twice; electrolytically charged for 1 month in 0.1M LiOD then frozen in liquid nitrogen, dipped in concentrated Na₂S solution and transferred to n-counting cell.ⁿ Four 1–2-mm-thick discs of different types of Pd (Johnson-Matthey) spot welded to Pd wires plus strained Pd Wire.^o Electrolyte reused after a previous run with a Ti cathode and Pt anode.^p Material heated to 900 °C then quenched in water before use.^q AuCN in Jones electrolyte replaced with NaAuCl₄.^r Current on for 36 h then changed between 200 mA and 20 mA every hour.^s Current on for 36 h then changed between 300 mA and 30 mA every 5 min.^t Current changed slowly over range shown during electrolysis period as anode disintegrated.^u Current on for 36 h then cycled off for 10 h and on for 2 h for period of 270 h. Then 5,000 s at 380 mA cathodic and 4,000 s at 10 mA anodic for rest of period.^v Current cycled on/off every 2 h. Neutron yield is for 'on' cycle.^w The errors assigned (1 σ , calculated over full run duration) vary somewhat because in an attempt (mostly in the earlier stages of the programme) to cover as wide a range of cell configurations as possible, pairs of unpurified + unpurified and purified + unpurified runs were not always carried out, and consequently allowance has to be made for slight differences in overall cosmic-ray neutron detection efficiencies caused by slight differences between the two nominally identical cells.^x These data were obtained using a different data acquisition system to drive the shuttle and accumulate data, set up to look for neutron bursts¹⁰.^y Cells exchanged every 5 min by hand.^z All electrolytes used for γ -ray work were 50:50 H₂O:D₂O except * which were 41:59 H₂O:D₂O.^{aa} For the present measurements, the system was calibrated with a set of standard γ -ray sources and a ²³⁸Pu/¹³C (α , γ) source emitting 6,129-keV γ -rays. The cells were positioned such that their cathodes were as close as possible to the detector crystal, and the detection efficiencies were calculated by integrating previously measured²¹ point efficiency functions over the cathode volume. Peaks were searched for at 5,488, 4,977 and 4,466 keV corresponding to the expected location of the full energy, single-escape and double-escape peaks using the method recommended in ref. 22.

nuclear reactions proceeding predominantly non-radiatively to ^4He , four cathodes (Table 2, first four entries) were analysed by vacuum fusion/mass spectrometry. The high vapour pressure of palladium at the melting point caused difficulties, so internal standards were prepared by ion implantation of 10^{13} and 10^{15} atoms of ^4He into samples cut from the cathodes. Detection limits for ^3He and ^4He determined in this way were $\sim 8 \times 10^{10}$ atoms per sample— $(1-10) \times 10^{11}$ atom g^{-1} . We found no ^3He or ^4He . The expected level, if fusion had been occurring at the rate reported in ref. 1, was $\sim 10^{16}$ atom g^{-1} .

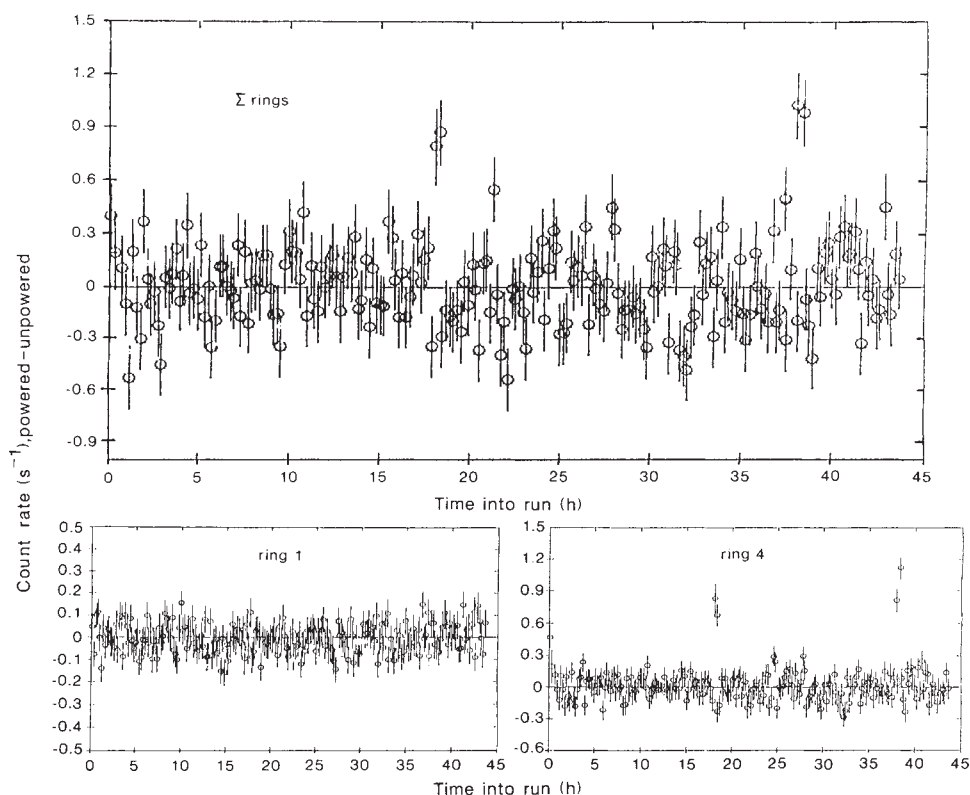
Evolution of hydrogen at a titanium cathode resulted, as is well known¹³, in a dense network of hydride precipitates, observable by standard metallographic methods, penetrating below the surface. In 0.1M D_2SO_4 electrolyte at 100 mA cm^{-2} , the network penetrated $\sim 30 \mu\text{m}$ in 1 h. Precious-metal deposition inhibited the electrolytic uptake, presumably by lowering the overvoltage and promoting gas evolution. Electrolysis in the 'brew' used by Jones *et al.*² resulted in hydride precipitates confined to the grain boundaries; other experiments showed that the presence of PdCl_2 in the electrolyte caused this effect, presumably as a consequence of palladium plating on the cathode. It is clear that expressions of the fusion rate that assume that the cathode has composition TiD_2 are completely misleading: metallography shows that the number of deuterium pairs is far fewer, hence the claimed fusion rate per deuterium pair is far higher than the figure given, and so the results are even more difficult to reconcile with expectations than had been implied².

Discussion

The interest in cold fusion has generated a large number of neutron counting experiments. It is well known that, in general, it is inadvisable to measure the signal + background and background of counting measurements at different times and in different physical locations (as in ref. 1), because of unexpected systematic variations. This is especially true for low-count-rate experiments. Compensating for variation of the background rate and assessing appropriate errors for the procedure chosen are particular problems. The work of Jones *et al.*² can be criticized for such errors¹⁸. It is notable that in ref. 2, only one run in fourteen showed a significant effect and then only because this particular run was assigned a smaller counting error than the others. Here we have attempted to minimize, by the shuttle procedure, uncertainty about background and counter variability and about error calculation. Given some of the more spectacular claims that have been made, we note that further caution is advisable because of the notorious sensitivity of $^{10}\text{BF}_3$ and ^3He proportional counters to humidity and of counter-amplifier systems to earth loops. In our work the neutron detectors were segmented, and the relationship between signals from the segments was well known, so spurious effects giving inconsistent signals from the segments could be identified.

Failure to reproduce the effects has been attributed by some to the need for rigorous exclusion of hydrogen and claims have been made that palladium electrodes must be cast and carefully degassed before use, to remove all traces of carbon and hydrogen impurity which might decorate dislocations or other high-energy

FIG. 3 Data from a typical run on the large high-efficiency neutron detector. These show the differences between count rates for powered and unpowered cells resulting from successive alternate shuttle positions for 40 g of titanium granules in 0.1M D_2SO_4 . The errors shown are 1σ errors for each five-minute counting period. In addition to results from the detector as a whole, results from two of the five rings of 2-inch diameter 107-cm active length $^{10}\text{BF}_3$ proportional counters are also shown. The counts in rings 2, 3 and 5 (not shown) were very similar to those in ring 1. The apparent bursts were seen only in ring 4 and are therefore a spurious effect. A tube (90-mm inner diameter) passes through the centre of the detector, and the shielding consists of 6 inches of borated resin, 1 mm of cadmium and 2 inches of lead. The neutron detection efficiency is high and is largely independent of neutron energy, varying from 48% for 0.5-MeV (Am/Li) neutrons to 40% for 4.2-MeV (Am/Be) neutrons. The background count rate is 4–5 count s^{-1} mostly neutrons from cosmic rays. Each of the five rings of counters has its own independent pre-amplifier, pulse-shaping amplifier and discriminator, and the mean energy of neutrons counted can be obtained from the ratio of counts in the outermost ring of counters (ring 5) to counts in the innermost (ring 1). Because neutron counts are distributed over five rings, and because of the 135- μs -mean time to capture, neutrons emitted simultaneously from a source as a burst are counted separately (as seen with the ^{252}Cf source, for example). The pre-amplifiers, high voltage components and insulators are all contained within a desiccated electrically screened box, and we eliminate earth loops



and induced electromagnetic pick-up ('aerial') effects from external coaxial-cable runs to the PDP-11/45 data-acquisition computer by using isolating high-frequency pulse transformers and by winding the coaxial cables many times around ferrite rings. The neutron detector and its immediately associated electronics are located in a temperature-controlled air-conditioned blockhouse with two-foot-thick concrete walls and roof.

sites (ref. 19 and R. A. Huggins, Workshop on Cold Fusion, Santa Fe, May 1989): this is however inconsistent with the postulate of a 'fusion' origin for the effect because, at low energies, p-d fusion is expected to be significantly faster than d-d fusion³. Furthermore, the original reports^{1,2} did not mention any special precautions to exclude atmospheric water vapour apart from careful covering of the electrolytic cells. We followed similar procedures, and found that degradation of the heavy water by exchange with atmospheric moisture occurred quite rapidly. Alternatively, it is claimed that it is essential to maintain a high current density for a considerable period (A. J. Appleby, Workshop on Cold Fusion, Santa Fe, May 1989) although it is not completely clear whether these latter claims are in fact reproductions of the effect reported in ref. 1 or are something different. In our neutron counting experiments, fresh dislocations were introduced by plastic deformation, some counting experiments were carried out at current density as high as 1 A cm^{-2} (Table 2) and in one experiment in the isothermal calorimeter a high current density was maintained for a considerable period (Table 1). Trace deposition of platinum on the cathode, supposedly causing a lowering of the overpotential for deuterium evolution and hence a lowering of the attainable deuterium level in the cathode, has also been suggested as an explanation for irreproducibility (M. Fleischmann, personal communication). We observed no effect when we used palladium anodes in our neutron counting experiments.

Timescales for hydrogen and deuterium loading consistent with the expected diffusion time (x^2/D where the diffusion coefficient, $D = 10^{-7} \text{ cm}^2 \text{ s}^{-1}$ and x is the radius or half-thickness of the specimen) were measured¹⁰ and nuclear counting and calorimetric experiments were always conducted over periods much longer than this. Furthermore, in the process of electrolytic loading of palladium, there will clearly be a concentration gradient, a moving phase boundary and the outer atomic layers of the metal will be saturated (possibly supersaturated) with deuterium¹⁷. It might be expected, therefore, that sufficiently sensitive equipment would detect any fusion process well before the material is completely loaded. Because our neutron detection sensitivity was $\sim 10^5$ – 10^6 times greater than that of Fleischmann *et al.*¹, it seems unlikely that any greatly enhanced fusion process associated with the absorption of deuterium into palladium, giving rise to neutron emission, is occurring. We are, of course, aware that it is always possible to construct essentially untestable theories involving hypothetical special conditions of the metal or of its surface. Careful characterization of materials for which positive results are claimed is therefore of great importance.

It has been argued that the neutron branch of the d-d reaction might be completely suppressed in favour of the (t+p) branch and it has been further argued (S. Pons, personal communication) that the tritium produced as a result of a nuclear process inside the electrode need not necessarily exchange with the electrolyte and might not therefore be detected. However, the other product of such a nuclear process, a high-energy proton, should be detectable by its interaction with the lattice, including neutron emission: we estimate, knowing the rate of energy loss of the protons and by comparison with (p,n) reaction cross-sections for neighbouring elements, a yield of $\sim 10^{-6}$ neutron per proton, implying a neutron yield in our counting experiments of as much as 10^4 s^{-1} if fusion at the rate reported in ref. 1 were to proceed entirely through the (t+p) branch.

In view of the rather large d-d separations in both PdD and TiD₂, it might be argued that fusion requires some non-equilibrium state in the lattice—perhaps at the α/β phase boundary in palladium or at the tips of the growing TiD₂ needles or at a lattice defect—where the d-d or p-d separation could be greatly reduced. Here we created non-equilibrium situations by pulsing the current but we detected no neutron emission (Fig. 4).

Some explanations of the apparent excess heat production have emphasized recombination processes at catalytic metal

surfaces²⁰: apart from occasional explosions, however, which did not result in any detectable neutron emission, we found, by comparison of the volume of water added to maintain the cell volume with the electrolysis charge passed, that this was not a significant effect (in agreement with others⁹). Recombination in the gas space above the liquid, either on exposed cathode surface or catalysed by colloidal metal particles eroded from the cathode, might however account for some of the observations of bursts of heat recently reported (M. Fleischmann and S. Pons, Electrochemical Society Meeting, Los Angeles, May 1989). In discussing the claims in ref. 1, we prefer to focus on characteristics of the 'simple' Fleischman, Pons and Hawkins (FPH) calorimeters, because we have only observed small effects (at the level of the inherent uncertainties), which might mistakenly be claimed as arising from cold fusion, in the one type of calorimeter (FPH type) that has major calibration difficulties. Cells using the same electrode and electrolyte materials operated in calorimeters that did not have these problems exhibited none of these effects.

There are two points regarding the calibration that could have a profound effect on the apparent results obtained with FPH-type calorimeters: the first concerns when the calibration is performed and the second how it is performed. Concerning the first point, it seems from our work and that of Lewis *et al.*⁸ that a calibration performed during the first 10,000 min of electrolysis could be seriously in error and lead to an erroneous conclusion that subsequently, rather than being in balance, the cells were exothermic. Concerning the second point, Fleischmann *et al.*¹ describe calibration using the internal-resistance heater, by measurement of Newton's-law-of-cooling losses. Typically, this procedure might involve the application of power to the heater while electrolysis was occurring, following the temperature-time trace until a steady state was obtained, then switching the heater off and following the cooling curve. This procedure gives an approximation to the differential calorimeter constant, $k_d = d(\Delta P)/d(\Delta T)$ (where P is power and T is temperature) at the operating temperature of the cell and can give rise to errors in two ways. First, because any calibration sequence would require

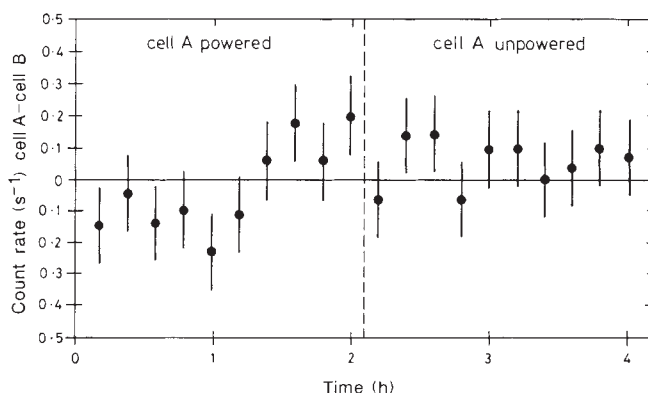


FIG. 4 Time dependence of neutron emission from an electrolytic cell with melt-spun Pd ribbon cathode and Pd foil anode, with power alternately turned on for deuterium loading and off for relaxation every 2 h by the data-acquisition computer: see Table 2 footnote 1. The total length of the run was 110 h, and the sums of the 'on' and 'off' data are shown. After 90 h of running, we added $1 \mu\text{M Pb}^{2+}$ (to poison the surface by Pb deposition) to the solution. There was no difference between the data in the presence and absence of Pb^{2+} , and so data for the whole run was combined. This run was undertaken to enhance the appearance of the non-equilibrium effects discussed in the text. It is assumed that the ribbon would have had a high density of grain boundaries, dislocations and other lattice defects. Because the 'on' and 'off' periods were several times the characteristic diffusion time for the ribbon, it is assumed that the composition was cycling in the β -phase region between the fully loaded condition ($D/\text{Pd} \approx 0.83$) and the limit of the ($\alpha + \beta$) phase field ($D/\text{Pd} \approx 0.65$).

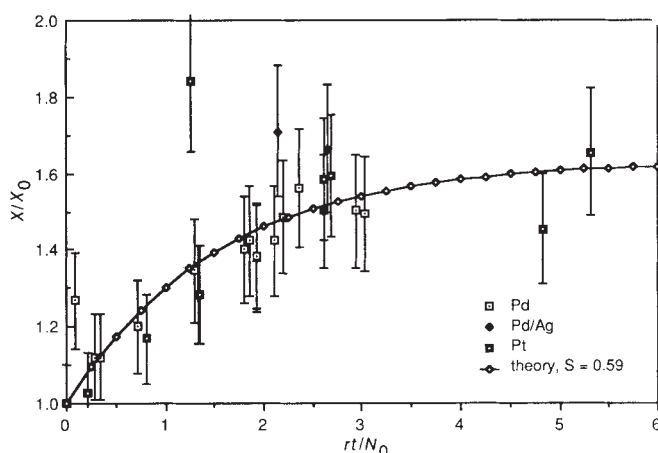


FIG. 5 Tritium enrichment by electrolysis in open cells at constant volume, with both Pd and Pt cathodes: relative count rate X/X_0 against amount of electrolysis rt/N_0 (symbols defined in the text). Initial count rate (efficiency corrected) was $13.0 \pm 0.6 \text{ kBq ml}^{-1}$. Errors are 1σ . The Pd cells were the FPH calorimeters, sampled after 490–660 h, the cells in the first section of Table 2 (footnotes a–e) sampled at the end of the run and one cell from the high-efficiency neutron detector—footnote k in Table 2. The line is the fit to equation (1).

5–10 h, extrapolation would be required to obtain the correct value at the reference liquid level: an estimated error of 20% or more could result. Furthermore, because the evaporation of the cell contents increases markedly with increasing temperature, the baseline slope would increase with increasing temperature and the effects of this sort of error would become correspondingly more marked: the claimed effects were indeed greatest in the cells run at the highest input power. Second, these calorimeters are nonlinear (equation (3), Fig. 2 legend), with the effect being significant when the temperature gradient is large: they cannot be described by a simple Newton's-law-of-cooling constant. If a differential calorimeter constant is used to derive the input power, the calculated output would be (from Fig. 2 legend, equation (3))

$$P_{\text{app}} = k_d \Delta T = (k_{\text{sb},0} + k_c) \Delta T + 3(k_{\text{sb},0}/T_0)(\Delta T)^2$$

so that, in comparison with the correct output (equation (3))

$$P_{\text{app}} - P = 3k_{\text{sb},0}(\Delta T)^2/2T_0$$

If the power applied to the heater is significant compared with the electrolysis power then this error will be even greater. Back

calculation from the data given in ref. 1 shows that the claimed effects were largest for Joule input powers on the order of 6 W. For our FPH-type calorimeters, this would have given $\Delta T \approx 50\text{--}60^\circ\text{C}$. Therefore, had these cells been calibrated in this way, an apparent heat excess on the order of 0.8–1 W would have been observed. Our method of calibration (see Fig. 1 caption) considerably reduced the effect of these sources of error. The original report¹ was clearly preliminary in nature, and it is evident from the above that the claims made therein cannot be assessed in the absence of a detailed description of the experimental procedure used and of the methods used to compensate for the systematic errors inherent in the use of a simple calorimeter.

We feel that our work has served to establish clear bounds for the non-observance of cold fusion in electrolysis cells, under carefully controlled and well understood experimental conditions and using well characterized materials. Further details are given in ref. 10. Claims of observations of cold fusion ought now to meet similar standards of data analysis and materials characterization so that a proper assessment can be made. □

Received 8 August; accepted 16 October 1989.

1. Fleischmann, M., Pons, S. & Hawkins, M. *J. electroanal. Chem.* **261**, 301–308 (1989); erratum **263**, 187 (1989).
2. Jones, S. E. *et al. Nature* **338**, 737–740 (1989).
3. Koonin, S. E. & Nauenberg, M. *Nature* **339**, 690–691 (1989).
4. Sun, Z. & Tomanek, D. *Phys. Rev. Lett.* **63**, 59–61 (1989).
5. Leggett, A. J. & Baym, G. *Nature* **340**, 45–46 (1989).
6. Gai, M. *et al. Nature* **340**, 29–34 (1989).
7. Zeigler, J. F. *et al. Phys. Rev. Lett.* **62**, 2929–2932 (1989).
8. Lewis, N. *et al. Nature* **340**, 525–530 (1989).
9. Cunneane, V. J., Scannell, R. A. & Schriffirin, D. J. *J. electroanal. Chem.* **269**, 163–174 (1989).
10. Williams, D. E. *et al. Harwell Report AERE R-13606* (HMSO, London, 1989).
11. Edwards, G., Findlay, D. J. S. & Lees, E. W. *Ann. Nucl. Energy* **9**, 127–135 (1989).
12. Lees, E. W., Patrick, B. H. & Bowey, E. M. *Nucl. Instrum. Meth.* **171**, 29–41 (1980).
13. Fouroulis, Z. A. *J. electrochem. Soc.* **128**, 219–221 (1981).
14. Östlund, H. G. & Werner, E. in *Tritium in the Physical and Biological Sciences* Vol. 1, 95–104 (Int. Atomic Energy Agency, Vienna, 1962).
15. Kaufmann, S. & Libby, W. F. *Phys. Rev.* **93**, 1337–1334 (1954).

16. von Buttlar, H., Vielstich, W. & Barth, H. *Ber. Bunsenges phys. Chem.* **67**, 650–657 (1963).
17. Smith, D. P. *Hydrogen in Metals* (University of Chicago Press, 1948).
18. Carpenter, J. M. *Nature* **338**, 711 (1989).
19. Gittus, J. & Bockris, J. *Nature* **339**, 105 (1989).
20. Kreysa, G., Marx, G. & Plieth, W. *J. electroanal. Chem.* **266**, 437–441 (1989).
21. Croft, S. *Nucl. Instrum. Meth. Phys. Res. A* **281**, 103 (1989).
22. Anicin, I. V. & Yap, C. T. *Nucl. Instrum. Meth. Phys. Res. A* **259**, 525–528 (1987).
23. Savitzky, A. & Golay, M. J. E. *Anal. Chem.* **36**, 1627–1639 (1964).

ACKNOWLEDGEMENTS. We thank Johnson-Matthey (Dr I. McGill and Dr M. Doyle) for assistance and for the loan of special samples of palladium. IMI Titanium (Mr J. R. B. Gilbert) provided titanium and information. We thank the following for assistance: P. Fozard, M. Newman, J. Monahan, R. Morrison, H. Bishop, V. Moore and colleagues, J. Asher, P. W. Swinden, A. M. Leatham, R. A. P. Wiltshire, D. A. Webb, J. O. W. Norris, R. Roberts, R. Crispin, C. Westcott, R. Neat, D. Robinson, H. Watson and G. McCracken. We also thank Dr D. Schriffirin, Dr P. T. Greenland and Dr D. Morrison for discussions and Dr R. Bullough for support. This work was financed by the Underlying Science programme of the UKAEA, and UK Department of Energy and the Commission of the European Communities.

Induction of self-tolerance in mature peripheral B lymphocytes

Christopher C. Goodnow, Jeffrey Crosbie, Helle Jorgensen, Robert A. Brink & Antony Basten

Centenary Institute of Cancer Medicine and Cell Biology, University of Sydney, New South Wales 2006, Australia

In transgenic mice, mature peripheral B lymphocytes in lymphoid follicles, like immature B cells, are rendered tolerant by encounter with self-antigen, provided receptor occupancy by self-antigen exceeds a critical threshold. The tolerant state of the B cell is closely correlated with down-regulation of membrane IgM but not IgD antigen-receptors. Identical changes in antigen-receptor expression occur in a subset of follicular B cells in nontransgenic mice, suggesting that clonally silenced self-reactive cells are common in the peripheral B-cell repertoire.

To maintain a state of immunological self-tolerance, potentially self-reactive lymphocytes must be prevented from mounting immune responses to autologous antigens. In the case of T lymphocytes, irreversible clonal deletion of self-reactive cells has recently been shown to play an important role in maintaining self-tolerance within the T-cell repertoire¹⁻⁵. Selective deletion of self-reactive T cells occurs within the thymus by a mechanism similar to that originally proposed by Lederberg⁶, whereby the interaction between self-antigen and the T-cell receptor at an immature stage in T-cell differentiation aborts any further development of the cell⁷. Such a censoring mechanism is effective in maintaining self-tolerance within the peripheral T-cell repertoire because the sequence and specificity of T-cell receptors remains constant after receptor gene rearrangement⁸⁻¹⁰. Self-tolerance in B-lymphocytes, however, poses an additional problem because the antigen-binding sites of immunoglobulins are further diversified after B cells have left the bone marrow, by hypermutation of immunoglobulin genes in mature, peripheral B cells^{11,12}. Hypermutation of immunoglobulin genes not only provides the extra diversity needed for production of high-affinity antibodies to most foreign antigens^{11,12}, but also results in generation of self-reactive B-cell variants at a detectable frequency^{13,14}. To prevent production of high-affinity pathogenic autoantibodies by this 'second wave' of self-reactive B cells, a mechanism of self-tolerance is required that can act at multiple stages of B-cell development.

High-affinity antibody responses to xenogeneic or allogeneic proteins or cells have long been known to be restricted to epitopes that differ between immunogen and self, despite the presence of many cross-reactive epitopes on such antigens^{15,16}, and this principle underlies the widespread practical use of species- or allele-specific antisera and monoclonal antibodies. As T-cell 'help' is potentially available to any B cell capable of binding and presenting the immunogen¹⁷, the restriction of the antibody response to those epitopes on xeno- or alloantigens that are not cross-reactive with self provides a clear indication that self-reactive B cells are censored at some stage during the generation of a high-affinity antibody response¹⁶. Indeed, evidence for induction of tolerance at multiple stages of B-cell development, including mature and memory B cells, has been repeatedly found in *in vivo* models of tolerance induction to

exogenous antigens administered in an appropriate form¹⁸⁻²⁶. It has been difficult, however, to determine the fate of the unresponsive mature peripheral B cells in these models because of the low-frequency of antigen-specific B cells. Also, the possibility remains that some examples of *in vivo* B-cell unresponsiveness may involve immunoregulatory mechanisms not directly related to self-tolerance, such as exhaustive differentiation into plasma cells, feedback effects due to antigen-antibody complexes and receptor blockade^{20,21,26}.

To increase the frequency of potentially self-reactive B cells, thereby allowing their fate to be traced *in vivo*, we have recently developed a transgenic mouse model in which hen-egg lysozyme transgenic mice were mated with anti-lysozyme immunoglobulin-transgenic mice to create 'double-transgenic' offspring²⁷. The anti-lysozyme B cells from these mice encountered lysozyme as a self antigen throughout their development within the bone marrow and in the periphery. Despite this, the B cells were not deleted, but persisted in a functionally silenced state characterized by expression of high levels of membrane IgD but a 10-20 fold reduction in membrane IgM²⁷. The current paper describes the use of the same transgenic model to follow the fate of self-reactive B cells when they encounter self-antigen only at the mature stage of development, in peripheral lymphoid organs. As observed previously in conventional *in vivo* models, mature follicular B cells in the transgenic mice were fully susceptible to induction of self-tolerance. Also, the silenced state appeared identical to tolerance induced by encounter with self-antigen early in B-cell development, in that it required comparable levels of receptor occupancy, and was intimately linked to downregulation of membrane IgM but not IgD. A significant proportion of IgD⁺ follicular B cells in nontransgenic mice show an identical change in membrane IgM, indicating that clonally silenced self-reactive B cells may be frequent in peripheral lymphoid organs.

Low-lysozyme double-transgenic mice

In previous experiments²⁷, functional silencing of anti-lysozyme B cells took place in double-transgenic mice derived from matings of two parental lines. The first line, ML5, was the highest expressing of a series of transgenic lines bearing the hen-egg lysozyme (HEL) gene linked to the mouse metallothionein promoter. The other parental line, MD3, was one of six lines bearing rearranged immunoglobulin heavy and light chain genes encoding high affinity anti-lysozyme IgM and IgD, the two isotypes being expressed by differential RNA-splicing from the V- μ - δ heavy chain transgene. Lysozyme-binding B cells persisted in peripheral lymphoid organs of the resulting double-transgenic progeny, but they no longer secreted anti-lysozyme antibody and exhibited an altered phenotype, characterized by a 10-20 fold reduction in membrane IgM with no change in membrane IgD. Developing B cells in the bone marrow of the double-transgenic mice encountered lysozyme shortly after expressing surface IgM, as bound lysozyme could be demonstrated on these cells (see below) and because down-regulation of membrane IgM was observed even on newly differentiated B220^{low} bone marrow B cells (D. Y. Mason *et al.*, manuscript in preparation).

In terms of B-cell receptor occupancy in the double-transgenic mice, the mean concentration of lysozyme in the serum of the

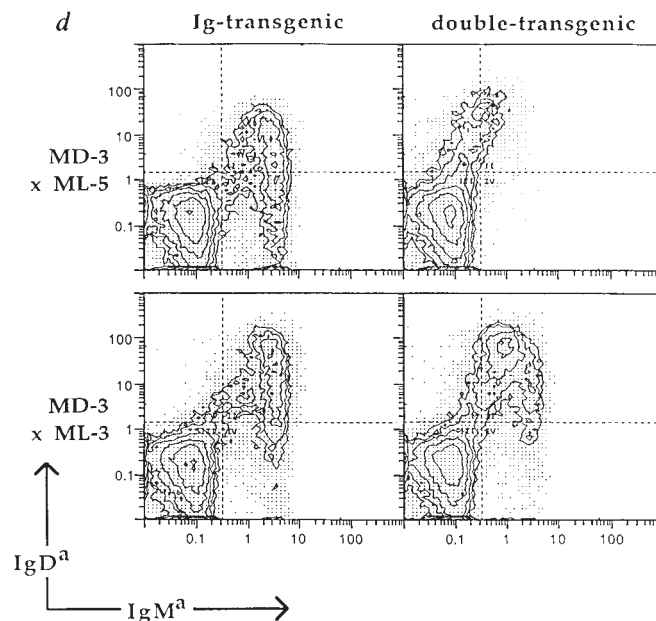
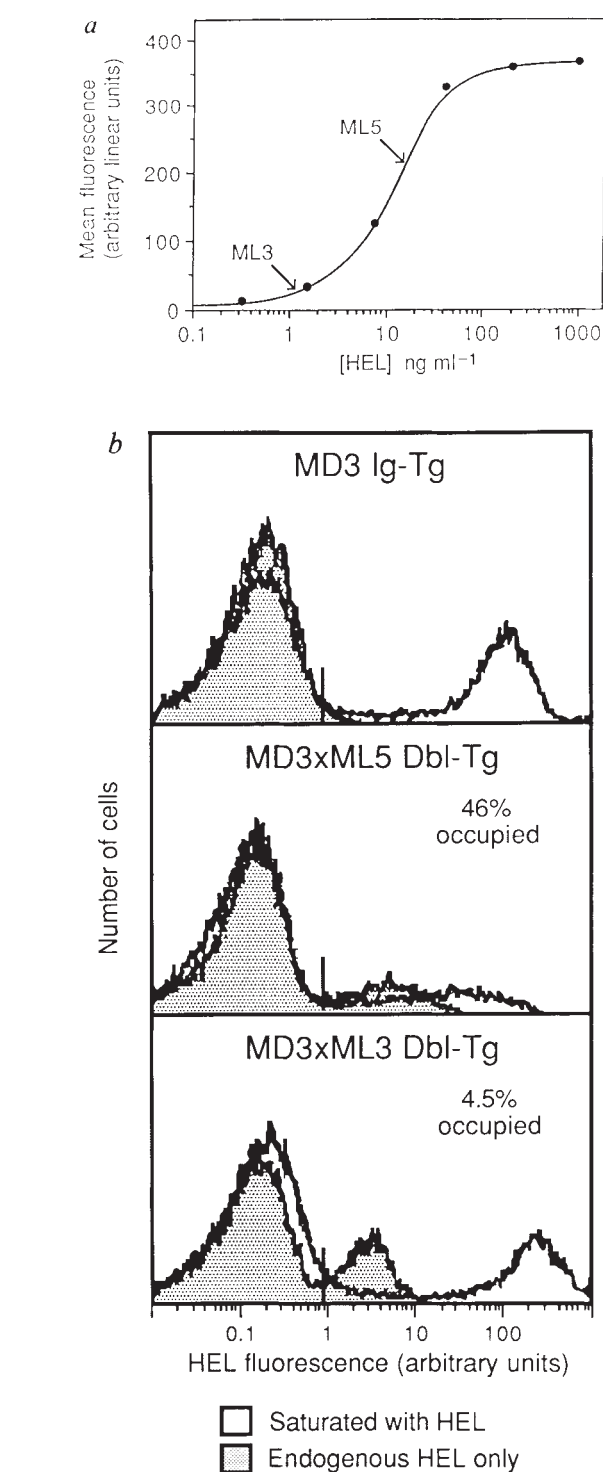
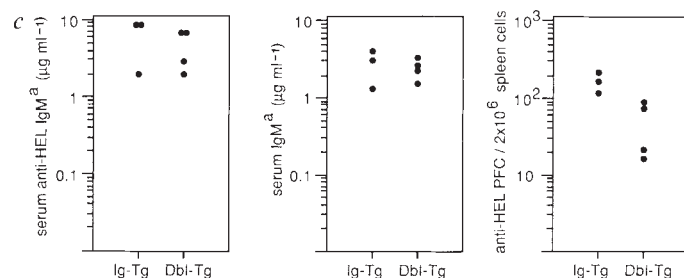


FIG. 1 Lysozyme-binding B cells in ML3 × MD3 double-transgenic mice exhibit decreased antigen receptor occupancy (*a, b*), failure of B-cell silencing (*c*), and absence of IgM-downregulation (*d*). *a*, Binding of lysozyme to MD3 immunoglobulin-transgenic spleen cells *in vitro* titrated over a range of lysozyme concentrations, by measuring subsequent binding of a biotinylated second anti-lysozyme monoclonal antibody on the FACS. *b*, Both immunoglobulin- and double-transgenic spleen cells exposed to saturating concentrations of lysozyme (unshaded histograms) or no lysozyme (shaded histograms) *in vitro* before second-step staining with the biotinylated anti-lysozyme monoclonal antibody. Positively staining cells in the shaded histograms therefore reflect cells which have bound lysozyme *in vivo* before preparation and staining *in vitro*. *c*, Spontaneous secretion of transgene-encoded (a-allotype) anti-lysozyme IgM measured in the serum of MD3 × ML3 double-transgenic mice and immunoglobulin-transgenic littermates by lysozyme-binding enzyme-linked immunosorbent assay (ELISA) (*c*, left panel) or by IgM^a allotype ELISA (*c*, centre), and by quantitating the number of anti-lysozyme IgM plaque-forming cells (PFC) in the spleen (*c*, right). Each point represents the value from an individual mouse. *d*, Expression of transgene-encoded membrane IgM and IgD on spleen cells from MD3 × ML5 or MD3 × ML3 double-transgenic mice and immunoglobulin-transgenic littermates determined by two-colour FACS analysis. Dots and contours are set respectively at 1 or more cells and 5, 10, 20, 40, 80, and 160 cells per 4 × 4 channel grid.

METHODS. Transgenic mice were derived from C57BL/6 founders as described²⁷ and maintained by backcrossing to C57BL/6. Lysozyme-binding was measured by incubating spleen or bone marrow cells with either no HEL or the indicated concentrations of HEL on ice for 25 min, followed by staining with biotinylated HyHEL9 and streptavidin-phycoerythrin as described previously²⁷. Logarithmic amplifiers on the FACS were calibrated using a series of standards, and the resulting nomogram used to convert mean fluorescence channel number to mean linear fluorescence. Per cent receptor occupancy was then derived by determining the mean fluorescence of the positively staining cells, using the gates shown, in the absence of any added HEL (shaded histograms), and dividing this value by the mean fluorescence of the same population of cells when all available receptors were saturated with 200 ng ml⁻¹ HEL (unshaded histograms). Anti-lysozyme IgM^a was measured at several dilutions from sera by ELISA on HEL-coated 96 well plates as previously, except that the assay was developed with biotinylated anti-IgM^a monoclonal antibody RS3.1 (ref. 51) followed by avidin-alkaline phosphatase (Sigma). Absolute concentrations were determined relative to a standard curve of culture supernatant from SP2/0 cells transfected with the anti-HEL immunoglobulin genes. Measurement of serum IgM^a allotype, anti-HEL direct plaque-forming cells, and two-colour FACS analysis were all as described previously²⁷.



parental ML5 mice (17.4 ng ml^{-1}), and in ML5 $\times$ MD3 double-transgenic progeny, was close to that required to half-saturate the anti-lysozyme immunoglobulin receptors expressed by MD3 immunoglobulin-transgenic B-cells *in vitro* (Fig. 1a), and *in vivo* (Fig. 1b). Because the lysozyme concentration in ML5 $\times$ MD3 double-transgenic mice lies in the steepest portion of the receptor titration curve (Fig. 1a), a relatively small reduction in lysozyme expression would be expected to result in a large decrease in receptor occupancy *in vivo*. To examine the effect of such a decrease in receptor occupancy on B-cell silencing, the same line of immunoglobulin-transgenic mice, MD3, was mated with a different line of metallothionein-lysozyme transgenic mice, ML3, which expressed an order of magnitude less lysozyme (mean serum HEL, 1.4 ng ml^{-1}).

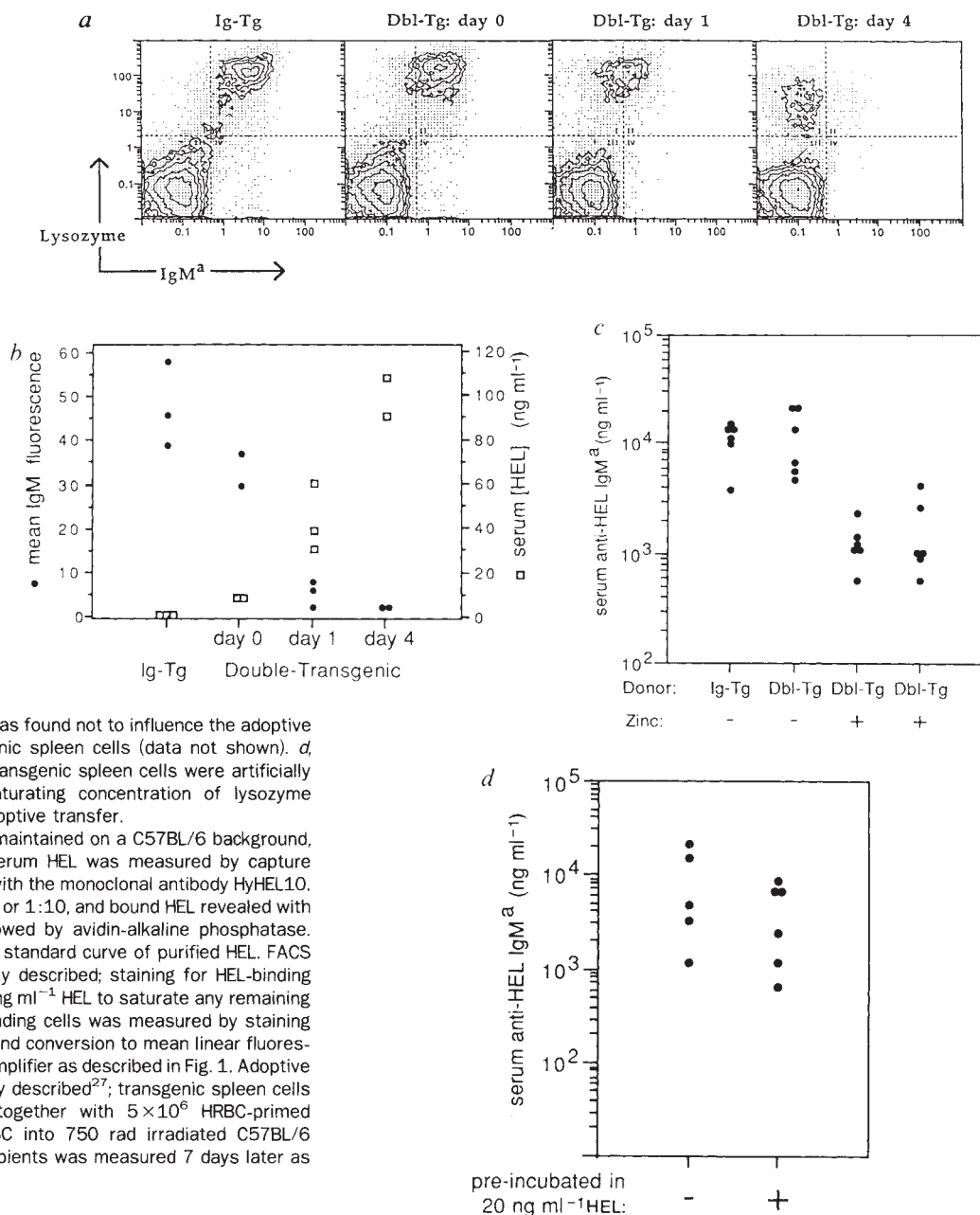
In ML3 $\times$ MD3 double-transgenic mice, only 4.5% of anti-lysozyme membrane immunoglobulin on splenic B-cells was occupied with lysozyme (Fig. 1b), compared with 46% occupancy in ML5 $\times$ MD3 double-transgenics. A similar difference in receptor occupancy was found on bone marrow B cells (data not shown), indicating that in both types of double-

transgenic mice, lysozyme was encountered early in B-cell differentiation. Comparison of these low-lysozyme double-transgenics with immunoglobulin-transgenic littermates revealed that, in contrast to the profound silencing of anti-lysozyme B cells observed in ML5 $\times$ MD3 double-transgenic mice²⁷, there was little or no reduction in the concentration of transgene-encoded anti-lysozyme IgM in the serum of ML3 $\times$ MD3 double-transgenics, and the number of cells secreting anti-lysozyme IgM in the spleen was only slightly lower than that of immunoglobulin transgenic littermates (Fig. 1c). In addition to the failure of tolerance in ML3 $\times$ MD3 double-transgenic mice, FACS analysis of their spleen cells revealed minimal down-regulation of membrane IgM on the anti-lysozyme B cells (less than twofold), compared with the 10- to 20-fold reduction in surface IgM on silenced B cells from ML5 $\times$ MD3 double-transgenic mice (Fig. 1d). Conversely, the T-helper cell compartment in ML3 transgenic mice was tolerant to lysozyme (S. Adelstein, unpublished results; see below), which is consistent with previous data showing a lower antigen concentration threshold for induction of unresponsiveness in T cells¹⁸.

FIG. 2 Zinc-induction of serum lysozyme in adult MD3 $\times$ ML3 double-transgenic mice is accompanied by rapid down-regulation of membrane IgM and induction of tolerance. MD3 $\times$ ML3 double-transgenic mice or immunoglobulin-transgenic littermates were given drinking water containing 25 mM zinc sulphate for 0, 1 or 4 days. All mice were killed at the same time. a, Lysozyme-binding versus IgM^a profiles of spleen cells obtained by two-colour FACS analysis b, Mean IgM levels on lysozyme-binding B cells expressed in arbitrary linear units compared to serum lysozyme concentration in individual animals at each time point c, Antibody responses of splenic B cells from individual immunoglobulin-transgenic, uninduced double-transgenic or 4-day induced double-transgenic mice measured in adoptive transfer; the dots represent the concentration of serum anti-HEL IgM^a in individual recipients of spleen cells from two different zinc-induced double-transgenic mice.

In other experiments, zinc exposure was found not to influence the adoptive response of immunoglobulin-transgenic spleen cells (data not shown). d, Receptors on MD3 immunoglobulin-transgenic spleen cells were artificially 'blocked' by exposure to a half-saturating concentration of lysozyme (20 ng ml^{-1}) on ice for 1 h before adoptive transfer.

METHODS. All transgenic mice were maintained on a C57BL/6 background, and used at 8–14 weeks of age. Serum HEL was measured by capture ELISA⁵² using 96-well plates coated with the monoclonal antibody HyHEL10. Serum was tested at a dilution of 1:5 or 1:10, and bound HEL revealed with biotinylated monoclonal HyHEL9 followed by avidin-alkaline phosphatase. Absolute values were derived from a standard curve of purified HEL. FACS analysis was performed as previously described; staining for HEL-binding included a first step exposure to 200 ng ml^{-1} HEL to saturate any remaining free receptors. Mean IgM^a on HEL-binding cells was measured by staining with fluoresceinated RS3.1 antibody and conversion to mean linear fluorescence using a calibrated logarithmic amplifier as described in Fig. 1. Adoptive transfer was carried out as previously described²⁷; transgenic spleen cells (10^4) were injected intravenously together with 5×10^6 HRBC-primed C57BL/6 cells and 2×10^8 HEL-HRBC into 750 rad irradiated C57BL/6 recipients. Serum anti-HEL in the recipients was measured 7 days later as described in Fig. 1.



Induction of lysozyme in adult mice

The findings in the low-lysozyme double-transgenic mice suggested that tolerance induction and the accompanying down-regulation in surface IgM depends on exposure of anti-lysozyme B cells to a threshold concentration of lysozyme which corresponds to a receptor occupancy of between 4.5% and 45% of surface immunoglobulin. They do not indicate, however, whether such a tolerance threshold is unique to immature B cells²⁶, nor whether mature B cells would in fact be triggered rather than silenced if above-threshold concentrations of lysozyme were encountered only after nontolerant B cells had matured in adult animals. To answer these questions, the serum lysozyme concentration in ML3 × MD3 double-transgenic mice was elevated above the putative threshold by exploiting the heavy metal inducibility of the metallothionein promoter²⁸. Adult ML3 × MD3 low-lysozyme double-transgenic mice were given drinking water containing 25 mM zinc for 0, 1 or 4 days. As a result, the serum lysozyme concentration was elevated to 29–60 ng ml⁻¹ in mice exposed to zinc for 1 day, and reached 89–108 ng ml⁻¹ following exposure for 4 days (Fig. 2b). In parallel with the rise in lysozyme concentration, membrane IgM on HEL-binding spleen cells from these mice was partially downregulated by day 1 (Figs 2a,b), and by day 4 of zinc exposure had fallen to a level comparable with that observed in the ML5 × MD3 double-transgenic mice. Importantly, at the intermediate timepoint there was no evidence for a bimodal distribution of membrane IgM levels expressed by the B cells (Fig. 2a) and at no stage was there a reduction in the number of B cells, excluding the possibility, therefore, that pre-existing IgM^{hi} cells were simply being replaced by a discrete population of IgM^{low} cells.

The functional status of anti-lysozyme B cells from the zinc-induced double-transgenic mice was assessed by adoptive transfer (Fig. 2c). In contrast to B cells from immunoglobulin-transgenic or uninduced ML3 × MD3 double-transgenic mice, B cells

from the mice exposed to zinc for 4 days responded poorly when challenged with lysozyme coupled to a foreign carrier (horse red blood cells, HRBC) following adoptive transfer into irradiated nontransgenic recipients together with HRBC-primed helper cells. In other words, the zinc-induced rise in serum lysozyme concentration and B-cell receptor occupancy led to changes in receptor phenotype and function in mature, peripheral B lymphocytes from adult ML3 × MD3 animals which were identical to those described previously in silenced B cells from ML5 × MD3 double-transgenic mice. The functionally silenced state of the B-cells cannot be accounted for simply by 'blocking' of antigen receptors with lysozyme for two reasons: first, less than 50% of the receptors were occupied in ML5 × MD3 double-transgenic mice (Fig. 1b), and second, artificially 'blocking' half the receptors with lysozyme *in vitro* before adoptive transfer failed to alter the responsiveness of immunoglobulin-transgenic cells (Fig. 2d).

Transfer of nontolerant B cells

The apparent susceptibility of mature peripheral B cells to IgM receptor down-regulation and silencing was examined in a second way, by transferring nontolerant spleen cells from immunoglobulin-transgenic mice into irradiated ML5 (high lysozyme) transgenic recipients or nontransgenic littermates (Fig. 3). The possibility of a donor T-cell response to recipient lysozyme was minimized by using C57BL/6 strain donors and recipients, which are low responders to lysozyme²⁹. Between 5 and 10% of the injected lysozyme-binding B cells homed to the spleen of both types of recipients. In the case of lysozyme-transgenic recipients, the cells exhibited a 10-fold down-regulation of surface IgM as early as 20 h after transfer (Fig. 3a) and persisted in this state for at least 10 days (data not shown). The functional capacity of the transferred B cells was examined by co-injecting the immunoglobulin-transgenic spleen cells with spleen cells from nontransgenic donors primed to HRBC to act

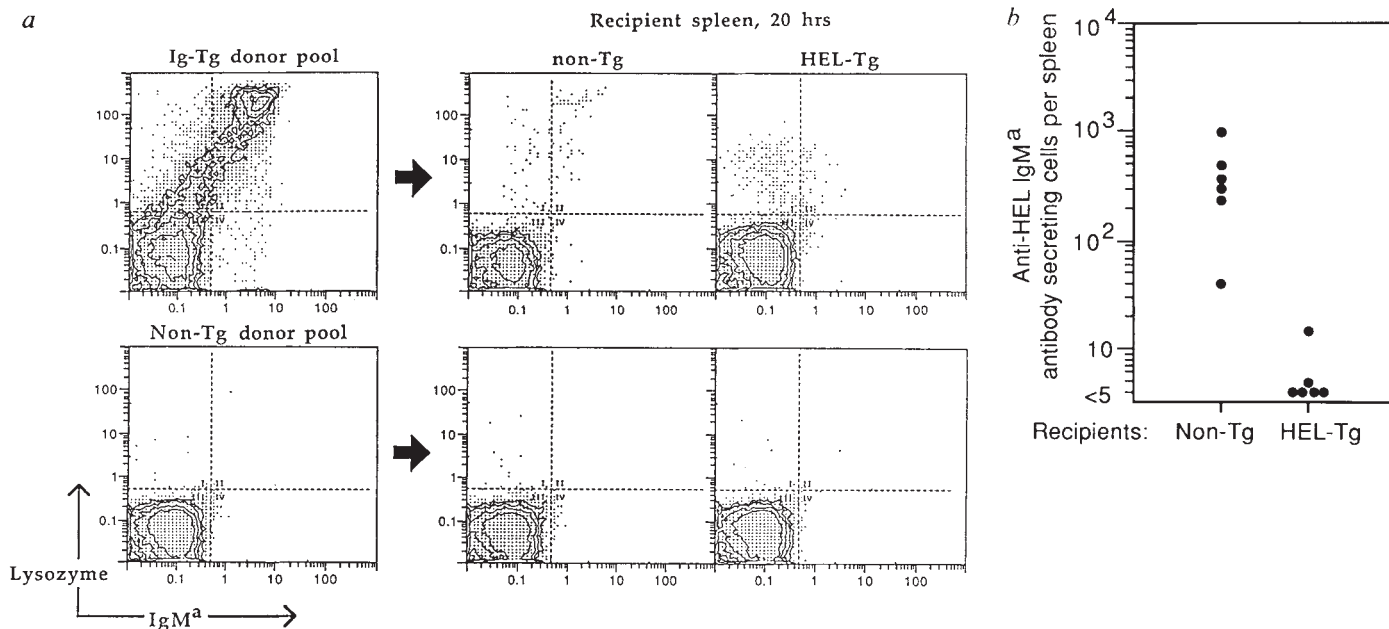


FIG. 3 Transfer of immunoglobulin-transgenic spleen cells into ML5 HEL-transgenic recipients leads to rapid downregulation of membrane IgM and induction of tolerance. *a*, For FACS analysis, 10^7 spleen cells from adult MD3 immunoglobulin-transgenic mice or nontransgenic controls were injected into 750-rad-irradiated ML5 transgenic mice or nontransgenic littermates, and recipient spleen cells analysed by two-colour FACS for HEL-binding against IgM at 20 h (*a*), 2.5 days or 10 days after transfer. Duplicate recipients at each timepoint gave similar results to *a*. *b*, For antibody responses, 10^6 immunoglobulin-transgenic cells were co-injected with 5×10^6 HRBC-primed C57BL/6 spleen cells into irradiated ML5-transgenic mice

or nontransgenic littermates, left for two days, and then challenged with 2×10^8 HEL-HRBC intravenously. Antibody responses in individual recipients were measured as anti-HEL IgM^a antibody-secreting cells in the spleen by spot-ELISA⁵⁰ five days after challenge.

METHODS. All donors and recipients were adult animals of C57BL/6 genetic background. Antibody secreting cells were measured by spot ELISA⁵³ in 24-well plastic trays coated with 1 mg ml⁻¹ HEL. Spots of bound antibody were revealed with biotinylated anti-IgM^a antibody RS3.1 followed by avidin-alkaline phosphatase.

as a source of helper T cells. Two days after cell transfer, the recipient mice were challenged with lysozyme coupled to HRBC (HEL-HRBC). In contrast to the good response obtained after 'parking' in nontransgenic hosts, immunoglobulin-transgenic B cells parked in HEL-transgenic recipients failed to mount an antibody response, as measured either by antibody-secreting cells in the spleen (Fig. 3b) or by serum antibody titres (data not shown).

Tolerance and IgM-down-regulation

The results from the low-lysozyme double-transgenic mice, the effects of zinc induction and the cell transfer experiments demonstrate a good correlation between selective down-regulation of membrane IgM but not IgD and clonal silencing of the B cell. The lack of down-regulation of membrane IgD cannot be explained in terms of failure of IgD to bind lysozyme as B cells from lines of immunoglobulin-transgenic mice expressing anti-lysozyme IgD alone bind lysozyme as efficiently as B cells from immunoglobulin-transgenic mice expressing IgM-alone, and B cells from double-transgenic mice expressing IgM alone show a more marked decrease in lysozyme binding which is directly proportional to the down-regulation of membrane IgM (unpublished observations). Thus the 4–10 fold decrease in lysozyme binding that accompanies the 10–50 fold down-regulation of IgM in tolerant double-transgenic mice expressing IgM and IgD (Figs 2a, 4b) can best be explained by binding of lysozyme to unchanged levels of membrane IgD. The degree of reduction in lysozyme binding observed may reflect both the fact that a subpopulation of the B cells do not express IgD (see below), and that in the absence of any down-regulation, IgM may be expressed at higher density than IgD.

A close relationship between selective down-regulation of membrane IgM and B-cell tolerance was also found in a reciprocal experimental approach, which involved mating different lines of anti-lysozyme IgM+IgD immunoglobulin-transgenic mice to the original (high lysozyme) ML5 transgenic line. Four out of the five double-transgenic combinations yielded similar results to ML5 × MD3 (a detailed comparison of all the lines will be described elsewhere). Double-transgenic offspring from the fifth immunoglobulin-transgenic line, MD5, were exceptional in that they failed to exhibit silencing of anti-lysozyme B cells, and continued to secrete high levels of transgene-encoded anti-lysozyme IgM into the serum (Fig. 4a). FACS analysis of spleen cells from both MD5 immunoglobulin- and double-transgenic mice showed that the B cells express high levels of lysozyme-binding IgM and IgD, and that the failure

of silencing in the double-transgenic mice was associated with minimal down-regulation of transgene-encoded membrane IgM (Fig. 4b). The explanation for the lack of membrane IgM down-regulation and B cell silencing in the MD5 line is not known, but may be related to a failure of allelic exclusion between the transgene and the endogenous heavy chain genes, which also appears to be unique to this line.

IgM-down-regulation in nontransgenic mice

In spleens from adult nontransgenic mice, three phenotypic populations of B cells have been defined on the basis of IgM and IgD expression³⁰. B cells expressing little or no IgD and high levels of IgM comprise one subset, designated 'population III' (Fig. 5), which is primarily composed of marginal zone B cells in addition to small numbers of immature B cells and Ly1 B cells^{31,32}. The other two subsets, 'populations I and II' are both relatively long-lived, IgD-positive, recirculating cells found in the follicular mantle zone³¹ and are distinguished only on the basis of their membrane IgM levels which vary over a 100-fold range (Fig. 5). Low expression of membrane IgM on population I follicular B cells seems to be a post-translational phenomenon, involving increased surface IgM turnover and/or decreased transport to the cell surface³³, but the underlying explanation for such changes in receptor expression is unknown.

In the transgenic mice, interaction with a threshold level of self-antigen induced a selective downregulation of membrane IgM and converted B cells with a population II phenotype into population I type cells (Fig. 5). The apparent equivalence of the IgD-positive B cells from immunoglobulin- and double-transgenic mice with populations II and I respectively is borne out by the identical expression of non-immunoglobulin B-cell

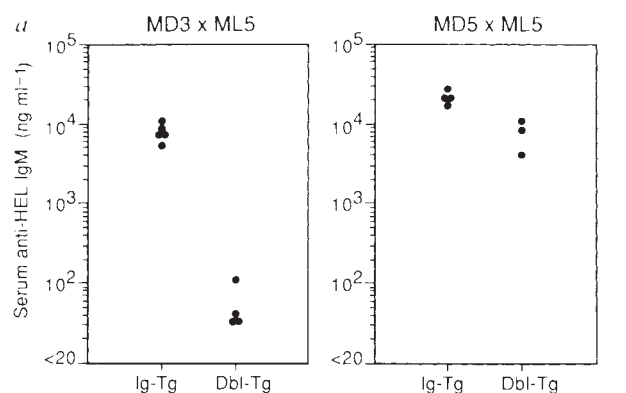


FIG. 4 Failure of tolerance and absence of downregulation of membrane IgM in MD5 × ML5 double-transgenic mice. *a*, Spontaneous secretion of anti-lysozyme IgM measured in the serum of MD3 × ML5 or MD5 × ML5 double-transgenic mice or immunoglobulin-transgenic littermates. *b*, Spleen cells from MD3 × ML5 or MD5 × ML5 double-transgenic mice or immunoglobulin-transgenic littermates analysed by FACS for expression of transgene-encoded IgM^a (top), or lysozyme-binding immunoglobulin (bottom) on B220-positive B cells. METHODS. Serum anti-HEL IgM antibody was measured as described in Fig. 1, except that sheep anti-mouse IgM-alkaline phosphatase was used as the developing reagent. FACS analysis was as described previously; IgM^a and B220 were detected with biotinylated RS3.1 and fluorescein-ated RA3-6B2 antibodies respectively.

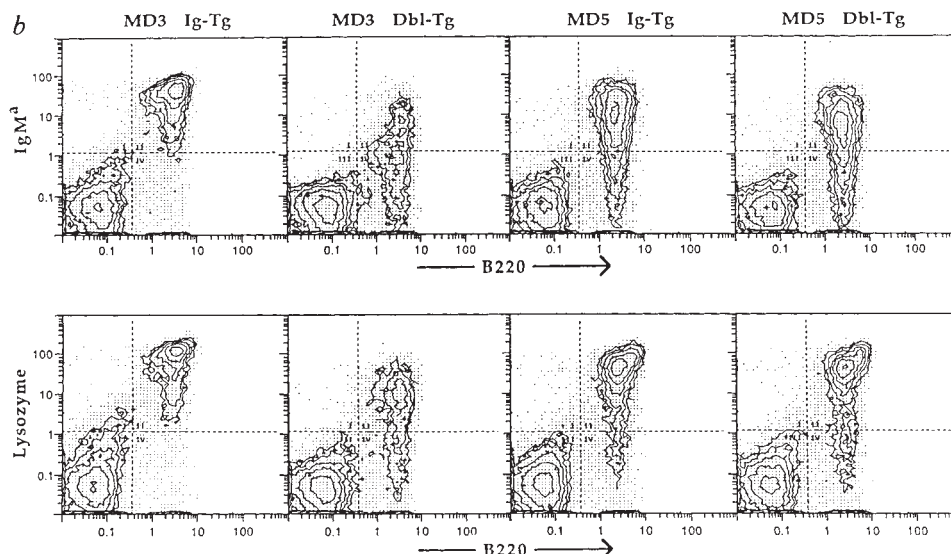
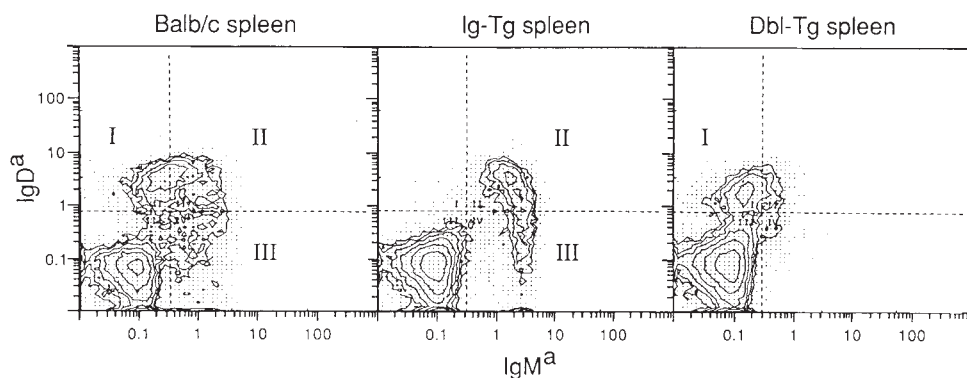


FIG. 5 IgD-positive splenic B cells from immunoglobulin and double-transgenic mice are phenotypically equivalent to B cell subpopulations II and I, respectively, in nontransgenic spleen. Spleen cells from MD4 immunoglobulin-transgenic or double-transgenic littermates and from an age- and sex-matched BALB/c control were stained for α -allotype IgD and IgM and analysed by FACS, as described in Fig. 1. I, II, and III indicate the three predominant B-cell subpopulations defined by relative expression of membrane IgM and IgD³⁰. To allow consistent use of anti-allotype (transgene-specific) reagents, BALB/c mice were used as non-transgenic controls in preference to C57BL/6 mice, as identical subpopulations are observed in both strains^{54,55}.



markers³⁴ and by immunohistological studies showing that the population II and I equivalents in the transgenic mice are indeed both located in the follicular mantle zone (D. Y. Mason *et al.*, in preparation). One interpretation of these findings is that population I and II follicular B cells normally represent a continuum of cells: at one extreme, some follicular B cells have high avidity for self antigens (lowest IgM), whereas at the other extreme, B cells exist that do not bind appreciably to any self antigen (highest IgM). The similarity may also hold in functional terms as population I cells with the lowest surface IgM, like double-transgenic B cells, have in fact been found to respond poorly in primary adoptive-transfer antibody responses to foreign antigens³⁵.

Discussion

Mature follicular B cells expressing transgene-encoded receptors were found to respond either positively to exogenous lysozyme presented on a foreign carrier, or negatively to endogenously synthesized lysozyme. These results support and extend earlier evidence for tolerance induction in mature B cells in nontransgenic mice¹⁸⁻²⁶, and raise the question of how self-reactive B cells are selectively tolerized. Because the stage of B-cell development does not seem to be critical for tolerance induction either in the transgenic model or in conventional *in vivo* models^{23,25}, some other variable must account for the immunogenicity of exogenous lysozyme compared with the tolerogenicity of autologous lysozyme. It is tempting to speculate that the factor responsible for determining whether the B cell becomes unresponsive or activated is the absence or presence of second signals derived from helper T cells during the initial contact with antigen, as originally proposed by Bretscher and Cohn³⁶. Thus, T cell 'help' would be absent during the initial encounter of B cells with above-threshold lysozyme concentrations in the zinc induction experiments, owing to the fact that the low levels of lysozyme expressed in uninduced ML3 mice are still sufficient to render the helper T cell compartment tolerant (S. Adelstein *et al.*, manuscript in preparation). Similarly, signals from helper T cells would be deficient in the parking experiments during the two days before challenge with HEL-HRBC because the C57BL/6 donor T cells, although not necessarily tolerant to recipient lysozyme, are unable to respond to lysozyme owing to their low-responder genotype²⁹. In both cases, subsequent encounter with HEL-HRBC would therefore have failed to elicit an antibody response, despite provision of optimal T-cell help, because the B cells had by then become intrinsically unresponsive.

The capacity of a threshold concentration of lysozyme corresponding to a receptor occupancy of between 4.5 and 46% to silence mature peripheral B cells provides a straightforward mechanism for preventing the generation of high affinity autoantibodies by hypermutation, and conversely, helps to explain the common occurrence of low affinity autoantibodies in normal

individuals. The demonstration of a B-cell subset with the phenotype (IgM^{low}, IgD^{hi}) of functionally silenced cells in the peripheral lymphoid tissue of nontransgenic mice lends weight to the possibility that such a mechanism may indeed play a major role in maintaining self-tolerance within the B-cell repertoire.

The close relationship between down-regulation of membrane IgM and induction of an unresponsive state in the B cell is reminiscent of ligand-induced receptor downregulation and desensitization in other biological systems³⁷⁻⁴⁰. It is unlikely that the decreased IgM receptor level itself accounts for the unresponsive state of the cell, as membrane IgD is still expressed at high levels and on its own is fully capable of mediating B-cell activation in immunoglobulin-transgenic mice (R. Brink *et al.*, in preparation) as well as in other models^{41,42}. Ligand-induced desensitization of IgM and IgD antigen receptors, in the absence of downregulation, has been shown to occur in mature B cells *in vitro*⁴³. By analogy with other signal-transduction systems^{37-40,44,45} biochemical changes affecting subsequent steps in the signalling pathway (reviewed in refs 42, 46, 47), such as phosphorylation or other alterations in membrane immunoglobulin-associated CD3-homologues⁴⁸⁻⁵⁰, could explain both the selective change in IgM-expression and the unresponsive state of the B cell. A detailed understanding of how tolerance mechanisms act on mature peripheral B cells may prove important for controlling pathological antibody responses in allergy and autoimmunity. □

Received 10 August; accepted 5 October 1989.

- Kappler, J. W., Roehm, N. & Marrack, P. *Cell* **49**, 273-280 (1987).
- MacDonald, H. R. *et al. Nature* **332**, 40-45 (1988).
- Kisielow, P. *et al. Nature* **333**, 742-746 (1988).
- Sha, W. C. *et al. Nature* **336**, 73-76 (1988).
- Pullen, A. M., Marrack, P. & Kappler, J. W. *Nature* **335**, 796-800 (1988).
- Lederberg, J. *Science* **129**, 1649-1653 (1959).
- Smith, C. A. *et al. Nature* **337**, 181-184 (1989).
- Chien, Y.-H. *et al. Nature* **309**, 322-326 (1984).
- Ikuta, K., Ogura, T., Shimizu, A. & Honjo, T. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7701-7705 (1985).
- Fink, P. J. *et al. Nature* **321**, 219-226 (1986).
- Berek, C. & Milstein, C. *Immunol. Rev.* **105**, 5-26 (1988).
- French, D. L., Laskov, R. & Scharff, M. D. *Science* **244**, 1152-1157 (1989).
- Diamond, B. & Scharff, M. D. *Proc. natn. Acad. Sci. U.S.A.* **81**, 5841-5844 (1984).
- Shlomchik, M. J., Aucoin, A. H., Pisetsky, D. S. & Weigert, M. G. *Proc. natn. Acad. Sci. U.S.A.* **84**, 9150-9154 (1987).
- Benjamin, D. C. *et al. A. Rev. Immun.* **2**, 67-101 (1984).
- Goodnow, C. C. & Basten, A. *Sem. Immun.* (in the press).
- Lanzavecchia, A. *Nature* **314**, 537-539 (1985).
- Chiller, J. M., Habicht, G. S. & Weigle, W. O. *Science* **171**, 813-815 (1971).
- Katz, D. H., Hamaoka, T. & Benacerraf, B. *J. exp. Med.* **136**, 1404-1429 (1972).
- Weigle, W. O. *Adv. Immun.* **16**, 61-122 (1973).
- Howard, J. G. & Mitchison, N. A. *Progr. Allergy* **18**, 43-96 (1975).
- Desaymard, C. & Waldmann, H. *Nature* **264**, 780-782 (1976).
- Metcalf, E. S., Schrater, A. F. & Klinman, N. R. *Immunol. Rev.* **43**, 143-183 (1979).
- Scott, D. W., Venkataraman, M. & Jandinski, J. J. *Immunol. Rev.* **43**, 241-280 (1979).
- Parks, D. E., Nelson, P. A., Walker, S. M. & Weigle, W. O. *Ann. N.Y. Acad. Sci.* **392**, 210-222 (1982).
- Nossal, G. J. V. *A. Rev. Immun.* **1**, 33-62 (1983).
- Goodnow, C. C. *et al. Nature* **334**, 676-682 (1988).
- Palmiter, R. D. *et al. Nature* **300**, 611-615 (1982).
- Gammon, G. *et al. Immunol. Rev.* **98**, 53-73 (1987).
- Hardy, R. R., Hayakawa, K., Parks, D. R. & Herzenberg, L. A. *Nature* **306**, 270-272 (1983).

31. MacLennan, I. C. M. & Gray, D. *Immunol. Rev.* **91**, 61–85 (1986).
32. Herzenberg, L. A. *et al. Immunol. Rev.* **93**, 81–102 (1986).
33. Yuan, D. *J. Immunol.* **132**, 1566–1570 (1984).
34. Goodnow, C. C. *et al. Cold Spring Harbor Symp. quant. Biol.* (in the press).
35. Zan-Bar, I., Vitetta, E. S., Assisi, F. & Strober, S. *J. exp. Med.* **147**, 1374–1394 (1978).
36. Bretscher, P. & Cohn, M. *Science* **163**, 1042–1049 (1970).
37. Hart, J. & Roth, R. A. *J. biol. Chem.* **262**, 15341–(1987).
38. Gill, G. N. *et al. Cold Spring Harbor Symp. quant. Biol.* **53**, 467–476 (1988).
39. Williams, L. T., Escobedo, J. A., Keating, M. T. & Coughlin, S. R. *Cold Spring Harbor Symp. quant. Biol.* **53**, 455–465 (1988).
40. Sherr, C. J., Rettenmier, C. W. & Roussel, M. F. *Cold Spring Harbor Symp. quant. Biol.* **53**, 521–530 (1988).
41. Goroff, D. K., Stall, A., Mond, J. J. & Finkelman, F. D. *J. Immunol.* **136**, 2382–2392 (1986).
42. Cambier, J. C. *et al. Immunol. Rev.* **95**, 37–57 (1987).
43. Cambier, J. C. *et al. Proc. natl. Acad. Sci. U.S.A.* **85**, 6493–6497 (1988).
44. Lefkowitz, R. J. *et al. Cold Spring Harbor Symp. quant. Biol.* **53**, 507–514 (1988).
45. Stryer, L. *Cold Spring Harbor Symp. quant. Biol.* **53**, 283–294 (1988).
46. Mueller, D. L., Jenkins, M. K. & Schwartz, R. H. A. *Rev. Immunol.* **7**, 445–480 (1989).
47. Defranco, A. L., Gold, M. R. & Jakway, J. P. *Immunol. Rev.* **95**, 161–176 (1987).
48. Hombach, J. *et al. EMBO J.* **7**, 3451–3456 (1988).
49. Sakaguchi, N. *et al. EMBO J.* **7**, 3457–3464 (1988).
50. Reth, M. *Nature* **338**, 383–384 (1989).
51. Schüppel, R., Wilke, J. & Weiler, E. *Eur. J. Immunol.* **17**, 739–741 (1987).
52. Smith-Gill, S. J., Lavoie, T. B. & Mainhart, C. R. *J. Immunol.* **133**, 384–393 (1984).
53. Sedgwick, J. D. & Holt, P. G. *J. Immunol. Meth.* **57**, 301–311 (1983).
54. Hardy, R. R., Hayakawa, K., Haaijman, J. & Herzenberg, L. A. *Nature* **297**, 589–591 (1982).
55. Herzenberg, L. A. *et al. Nature* **329**, 71–73 (1987).

ACKNOWLEDGEMENTS. We thank Pat Gregory and Tracy Anderson for technical assistance, and Andrew Cheung for animal husbandry. C.C.G. is supported by a fellowship from The Medical Foundation, University of Sydney. The research was funded by the National Health and Medical Research Council and the Australian Research Council.

The guanosine binding site of the *Tetrahymena* ribozyme

François Michel*, Maya Hanna, Rachel Green, David P. Bartel & Jack W. Szostak

Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114, USA

The self-splicing Group I introns have a highly specific binding site for the substrate guanosine. Mutant versions of the *Tetrahymena* ribozyme have been used in combination with guanosine analogues to identify the nucleotide in the ribozyme that is primarily responsible for recognition of the guanine base.

ONE of the substrates of the class of ribozymes known as the Group I introns is the nucleoside guanosine; an understanding of how this substrate is bound is essential to an understanding of the mechanism of catalysis of these RNA enzymes. All reactions performed by Group I introns seem to involve the binding either of guanosine or an internal G residue. The first step of self-splicing requires the specific binding of free guanosine, which becomes covalently added to the 5' end of the intron after cleavage of the 5' intron–exon junction¹. Group I introns invariably end with a 3' terminal G residue (for a review, see ref. 2), which is required for efficient cleavage and ligation at the 3' intron–exon junction³. Reactions that mimic circularization of the excised intron also require guanosine⁴.

It is clear that the Group I introns contain a well-defined guanosine-binding site. Bass and Cech^{5,6} used guanosine analogues to show that both the ribose and guanine moieties of guanosine are specifically recognized by the *Tetrahymena* intron. Here we present evidence that a phylogenetically invariant guanine residue in the highly conserved stem P7 in the catalytic core of the intron is the part of the guanosine binding site that interacts with and recognizes the guanine base.

Base-paired stem constrains binding-site location

The site that binds the 3'-terminal G residue of the *Tetrahymena* ribozyme must be physically close to the site that binds the 3' intron–exon junction. We therefore first attempted to clarify the nature of the 3' intron–exon junction binding site. A part of the intron called the internal guide sequence (IGS) is known to bind the 5' exon; formation of a base-paired stem (P10) between the 3' exon and the IGS has been proposed to align the two exons for ligation (ref. 7, and Fig. 1). The P10 pairing is not present in all Group I introns, however, and can be disrupted

in the *Tetrahymena* intron without affecting splicing⁸. These data indicate that the 3' intron–exon junction could be recognized by redundant mechanisms. A probable candidate for such an additional interaction is the recently identified P9.0 pairing between the two nucleotides immediately preceding the 3' terminal G residue, and two nucleotides in the segment (J7-9) connecting base-paired stems P7 and P9 (F. Michel, P. Netter, M. Xu and D. Shub, manuscript submitted). The P9.0 pairing was identified by sequence comparison analysis, and is widespread although not universal in the Group I introns.

By mutationally disrupting the P9.0 and P10 putative pairings, both separately and together (Fig. 1 and Table 1), we found that both contribute to binding the 3' intron–exon junction in the *Tetrahymena* intron. None of these mutations had marked effects on the rate of 5' cleavage. Mutants with only the P9.0 or the P10 pairing disrupted had relatively minor effects on 3' splicing. By contrast, when both P9.0 and P10 pairing were disrupted, 3' splicing was strongly affected; these introns accumulate 5' exon and intron–3' exon during splicing. Restoring either the P9.0 or P10 pairing by compensatory mutations invariably alleviates the defect in 3' cleavage and ligation. Thus,

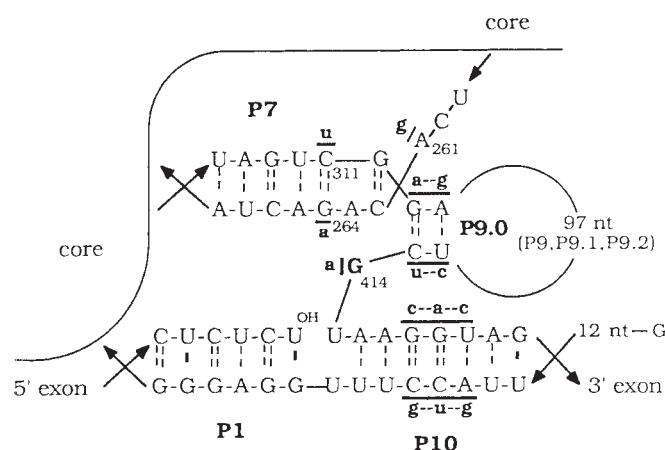


FIG. 1 Secondary structure elements surrounding the 3' terminal G (position 414) of the *Tetrahymena* intron. P1, P10, P7 and P9.0 are base-paired segments. (See refs 9–11 and 23 for nomenclature and a more complete secondary structure model of the intron.) The situation depicted is just before the ligation step. Mutations introduced in this work are indicated in bold, lower case letters next to the nucleotides that were substituted. nt, Nucleotide.

* Permanent address: Centre de Génétique Moléculaire du CNRS, 91190 Gif-sur-Yvette, France.

binding of the 3' intron-exon junction is ensured by base pairing of the nucleotides immediately flanking the terminal G residue to two distinct regions of the *Tetrahymena* intron. These two pairings strongly constrain the possible locations of the guanosine binding site.

Defective recognition of the 3' terminal G

We searched for an intron residue that might be involved in recognition of the 3' terminal G by screening mutants for the loss of ability to distinguish between a G and an A residue at the 3' end of the intron. Changing the 3' terminal G residue of the *Tetrahymena* intron (position 414) to a U or C residue results in normal cleavage at the 5' intron-exon junction, but greatly reduces the rates of 3' cleavage and ligation³. Under the same splicing conditions, we found that changing the 3' terminal G to an A residue led to delayed (but still accurate) 3' splicing.

Because all Group I introns end in a G residue we surmised that any nucleotide involved in recognition of this 3' terminal G was probably one of the very few residues that are conserved in all Group I introns. There are only two such invariant nucleotides close to P9.0 in the secondary structure of the *Tetrahymena* intron^{2,9-11}. These are residue A at position 261 (A261) in the joining segment J6/7, which we mutated to G (mutation A261 → G), and residue G264, which is part of the very highly conserved paired stem P7: the G264-C311 pair was mutated to A-U (mutation A-U).

Both the A261 → G and the A-U mutations greatly reduced the rate of 5' cleavage, but at high concentrations of GTP, cleavage was efficient enough to allow study of 3' splicing. In an A261 → G context, the rate of 3' splicing was further decreased when the 3' terminal G residue was mutated to A (Fig. 2a, c), showing that position 261 is not involved in recognition of the 3' terminal G. The A-U mutation, by contrast, greatly decreased the rate of 3' cleavage, but there was only a slight further decrease on changing the 3' terminal G residue to A (Fig. 2b, d). Either recognition of the 3' terminal nucleotide was no longer rate-limiting for the A-U introns, or they failed to distinguish G from A.

A guanosine analogue restores 5' cleavage

If the same binding site is used for both the 3' terminal G residue and the free guanosine used in 5' cleavage, the A-U mutant should have a greatly reduced affinity for free guanosine. Although we were not able to measure guanosine binding (dissociation constant, K_d) directly, we were able to measure the Michaelis constant K_m (depending on the precise reaction kinetics, K_m need not equal K_d), and we found that the K_m for 5' cleavage of the A-U intron with guanosine is 3.0 mM; this is two orders of magnitude higher than the K_m for 5' cleavage of

TABLE 1 Effect of disruption of the P9.0 and P10 pairings on exon ligation

P10	P9.0			
	Wild type	5'-disrupted	3'-disrupted	Restored
Wild type	0.91	0.42	0.60	0.85
5'-disrupted	0.74	0.05	0.011	0.20
3'-disrupted	0.92	0.10	0.31	0.88
Restored	0.96	0.59	0.73	0.94

Mutations introduced to disrupt the P9.0 and P10 pairings are shown in Fig. 1. Numbers shown are the fraction of 5' cleaved exon that has become ligated to 3' exon. Reactions were for 60 s; the fraction of reacted precursor varied from 22–87% depending on the mutant examined. Missplicing was minimal for all mutants, as the only band of significant intensity in between linear intron (~400 nucleotides (nt)) and 5' exon (57 nt) was the band generated by ligated exons (101 nt). The only exception was the mutant disrupted through mutation of the 3' branches of both P9.0 and P10: a band migrating to a position corresponding to 64 nt was observed that amounted to ~20% of the sum of ligated exons and 5' exon. Mutations were constructed by cassette mutagenesis as previously described¹⁹. Transcription, gel purification of [³²P]GTP-labelled full-length transcripts, splicing conditions (5 mM MgCl₂, 200 mM NH₄C₂H₃O₂, 20 mM Tris buffer pH 7.5, 0.02% SDS, 30 °C, 20-min preincubation, 2 mM GTP) and gel electrophoresis were essentially as described previously³. The radioactivity in electrophoretic bands was determined by direct scanning with a Betagen counter (Waltham, Massachusetts) and converted to molar amounts by correcting for the number of G residues.

the wild-type intron, which is 0.024 mM (Fig. 3a, Table 2). In addition, the maximum initial rate of the reaction is much reduced (0.0034 min⁻¹ instead of 0.29 min⁻¹); again, the catalytic constant K_{cat} need not reflect the chemical step of catalysis, but the simplest interpretation of these results is that the guanosine is not only poorly bound, but incorrectly orientated.

A model of possible hydrogen-bonding interactions between guanosine and the G264-G311 base pair is shown in Fig. 4. In this model, the G264 → A mutation would lead to a steric clash between two hydrogen-bond donors, resulting in both decreased affinity and incorrect positioning. To test this model, we synthesized the guanosine analogue 2-aminopurine ribonucleoside (2AP), which differs from guanosine in its distribution of hydrogen-bond donors and acceptors. The interactions proposed for the binding of 2AP to the A of an A-U base pair and for the binding of G to G-C are isosteric (Fig. 4). When we used 2AP in splicing reactions, we found that wild-type introns preferred G to 2AP, but A-U introns preferred 2AP to G (Table 2, Fig. 3b). The ability of the A-U mutant to perform 5' cleavage was fully restored at high concentrations of 2AP, and the K_m of the A-U mutant with 2AP (two hydrogen bonds; K_m = 0.43 mM) is

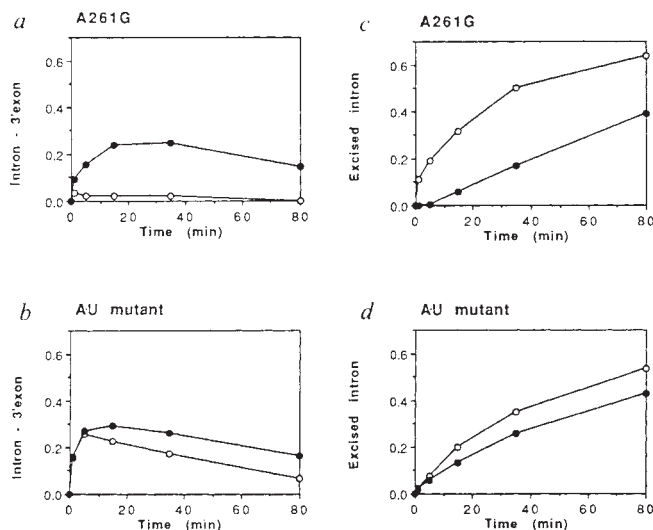


FIG. 2 Ability of two mutant versions of the intron to distinguish between 3' terminal G (○) and 3' terminal A (●). a, b, Accumulation of the intron-3' exon intermediate, expressed as fraction of total intron. a, A261G mutant. This intermediate does not accumulate with a 3' terminal G, but does with a 3' terminal A. b, AU mutant. This intermediate accumulates with similar kinetics whether or not the 3' terminal residue is A or G. c, d, Accumulation of excised intron after 3' cleavage. c, A261G mutant. Accumulation is delayed when the 3' terminal residue is A rather than G. d, AU mutant. Product accumulates with similar kinetics with either 3' terminal A or G. METHODS. Splicing was carried out in 15 mM MgCl₂, 40 mM Tris buffer pH 7.5, 2 mM spermidine, 5 mM DTT at 37 °C. Reactions were started by addition of GTP to a final concentration of 2 mM (<10% of precursor RNA had reacted after 80 min in controls without GTP).

lower than that of either the A·U mutant with guanosine (one hydrogen bond, steric clash; $K_m = 3.0$ mM) or the wild-type intron with 2AP (one hydrogen bond, no steric clash; $K_m = 1.1$ mM).

We used single mutants to show that the G residue of the G264·C311 base pair is responsible for discriminating between guanosine and 2AP. A G·C-to-G·U mutation still allowed some splicing with guanosine, but almost no splicing was detectable with 2AP (Fig. 3b). By contrast, a G·C to A·C mutation blocked splicing with guanosine, but allowed cleavage by 2AP.

The A·U mutation affects inhibition by arginine

The amino acid arginine inhibits self-splicing of the *Tetrahymena* intron by competing with guanosine¹². Yarus¹² proposed that the guanidino group of arginine can substitute for the two donor groups of guanine (Fig. 4). If so, splicing of the A·U mutant should not be inhibited by arginine, for the same reason that this mutant binds guanosine so poorly. We found that splicing of the wild-type intron was strikingly inhibited by 10 mM arginine, whereas no effect was seen on the activity of the A·U

mutant (Fig. 3c). Citrulline is an analogue of arginine with a substitution in the guanidino group which should enable it to bind to the A·U mutant intron (Fig. 4). Citrulline was reported to be only a very weak inhibitor of splicing¹³ and we observed only slight inhibition of splicing of the wild-type intron at 300 mM citrulline. But splicing of the A·U mutant was significantly more inhibited by citrulline (Fig. 3c), as predicted by the model for arginine binding.

2-aminopurine in exon ligation

If there is one binding site for both free guanosine and the 3' terminal G of the *Tetrahymena* intron, it should be possible to restore the A·U mutant to full 3' splicing activity by replacing the 3' terminal G residue by 2AP. Unfortunately, direct replacement of a nucleotide by an analogue in a long polynucleotide chain is not feasible at present. The exons to be ligated, however, need not be covalently linked to the intron and can be quite short¹⁴. We therefore synthesized two oligoribonucleotides that could be aligned on the IGS of the *Tetrahymena* intron. One oligonucleotide represented the 5' exon; the other represented

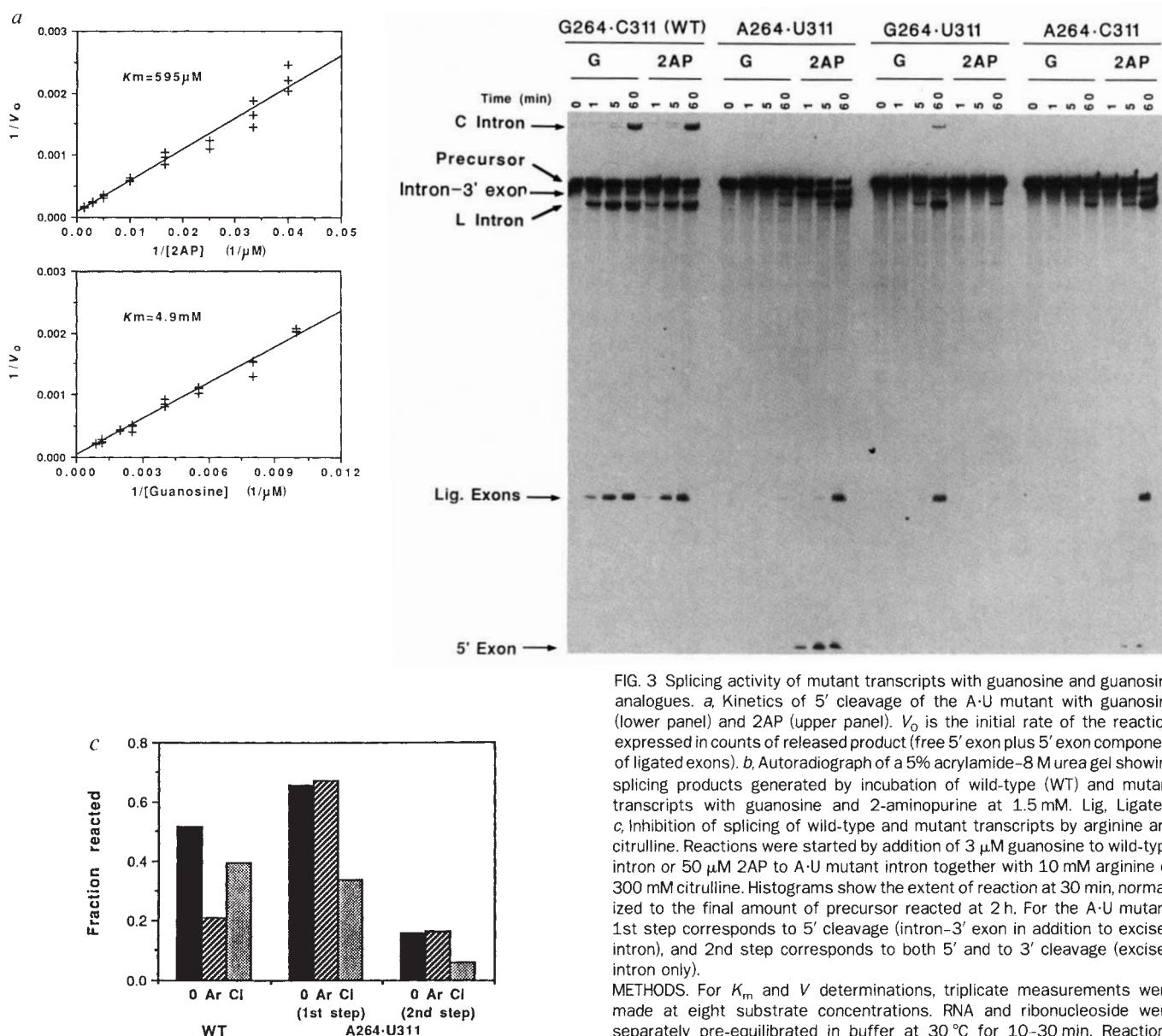


FIG. 3 Splicing activity of mutant transcripts with guanosine and guanosine analogues. *a*, Kinetics of 5' cleavage of the A·U mutant with guanosine (lower panel) and 2AP (upper panel). V_0 is the initial rate of the reaction, expressed in counts of released product (free 5' exon plus 5' exon component of ligated exons). *b*, Autoradiograph of a 5% acrylamide-8 M urea gel showing splicing products generated by incubation of wild-type (WT) and mutant transcripts with guanosine and 2-aminopurine at 1.5 mM. Lig, Ligated. *c*, Inhibition of splicing of wild-type and mutant transcripts by arginine and citrulline. Reactions were started by addition of 3 μM guanosine to wild-type intron or 50 μM 2AP to A·U mutant intron together with 10 mM arginine or 300 mM citrulline. Histograms show the extent of reaction at 30 min, normalized to the final amount of precursor reacted at 2 h. For the A·U mutant, 1st step corresponds to 5' cleavage (intron-3' exon in addition to excised intron), and 2nd step corresponds to both 5' and to 3' cleavage (excised intron only).

METHODS. For K_m and V determinations, triplicate measurements were made at eight substrate concentrations. RNA and ribonucleoside were separately pre-equilibrated in buffer at 30 °C for 10–30 min. Reactions were started by addition of ribonucleoside and stopped by addition of an equal volume of 90% formamide, 25 mM EDTA after 1 min (WT; mutant with 2AP) or 2 h (mutant with guanosine). RNA concentrations ranged from 45–90 nM. Other methods as in Table 1.

2AP respectively at the 3' splice site (ss). *a b* Principle (*a*)

METHODS. Truncated intron molecules lacking both 5' and 3'

5' exon	*	*	*	*	-	-	+
3' splice site (G)	-	-	-	+	*	*	*
Truncated intron (WT)	-	+	+	+	-	+	+
Anti-leader oligo	-	-	+	+	-	+	+

Wild-type		A264-U3II		Truncated intron
G--	2AP--	G--	2AP--	3' splice site
0	0	0	0	

the 3' splice site. The 3' splice site oligonucleotide began with a G residue; this G is equivalent to the 3'-terminal G of the normal intron, and is cleaved off during the ligation reaction. The two oligonucleotides were efficiently ligated by a shortened form of the *Tetrahymena* intron lacking both normal splice sites (Fig. 5). We took advantage of the fact that this reaction will run in reverse to generate a modified 3'-exon oligonucleotide beginning with 2AP: an oligonucleotide representing ligated exons was cleaved by the A·U mutant intron in the presence of free 2AP.

We could then compare the ability of the 3'-exon oligonucleotide beginning with G and its 2AP-substituted version to serve as 3' splice sites. As expected, the A·U mutant intron was unable to ligate a 3' exon beginning with G, but was able to ligate a 3' exon beginning with 2AP (Fig. 5c). By contrast, the 2AP oligonucleotide was not used by the wild-type intron. Therefore, in a reaction that mimics 3' cleavage and exon ligation, the G·C and A·U versions of the *Tetrahymena* intron show the same preferences for guanosine and 2AP as they do in 5' cleavage. This, together with the fact that the A·U mutant distinguishes poorly between G and A at the 3' intron-exon junction, shows that free guanosine and the 3' terminal G are not only recognized by the same nucleotide of the intron but most probably bind it by means of the same type of G·G base pair.

Discussion

The guanosine-to-G·C base triple postulated in Fig. 4 readily accounts for data reported by Bass and Cech⁵ concerning the ability of the *Tetrahymena* intron to use guanosine analogues. Both 1-methylguanosine and *N*-2-methylguanosine are poor substrates for splicing, as expected from the involvement of positions 1 and 2 in hydrogen bonding. By contrast, positions 7 and 8, which are not involved, can be substituted without effect. We have attempted to determine the role of position 6 by comparing the abilities of 2AP and 2,6-diaminopurine ribonucleosides to participate in 5'-cleavage: if 06 of guanosine is involved in hydrogen bonding, replacing it by a donor group would presumably create a steric clash, and therefore be highly detrimental. Our experiments (Table 2), however, show that 2,6-diaminopurine is nearly as effective as 2AP at saturating concentrations, whether wild-type or mutant introns are involved. It is possible that only two hydrogen bonds are formed between the bound guanine base and the intron, with stacking interactions also contributing to the high affinity and highly specific recognition of guanosine. The specificity of ribonuclease T1 towards guanine nucleotides rests similarly on the base stacking with a tyrosine residue, in addition to the formation of two hydrogen bonds with the main chain of this protein enzyme¹⁵. Lower stacking energies due to the loss of 06 might account for the elevated K_m of the A·U mutant for 2AP (and the even higher K_m for 2,6-diaminopurine), compared with that of wild type for guanosine.

It has been argued that Group I introns possess two G-binding sites¹⁶. This suggestion stemmed from the G-exchange reactions in which a 3'-OH group of a 3' terminal G attacks the phosphate

TABLE 2 Splicing of wild-type and mutant introns with guanosine and analogues

RNA	Purine nucleoside	K_m (mM)	V (min ⁻¹)	V/K_m (M ⁻¹ min ⁻¹)
Wild-type	Guanosine	0.027	0.34	12,500
		0.021	0.24	11,500
	2-Aminopurine	1.09	0.17	156
	2,6-Diaminopurine	2.2	0.14	63
		1.4	0.17	119
Mutant	Guanosine	2.4	0.0019	0.81
		3.7	0.0049	1.33
	2-Aminopurine	0.59	0.68	1,140
		0.28	0.54	1,930
	2,6-Diaminopurine	3.6	0.26	73
		1.8	0.25	139

Velocity of enzyme-catalysed reaction at infinite concentration of substrate (V) and K_m values were calculated from plots as in Fig. 3a. V is expressed as the fraction of precursor RNA converted per min, corrected for the maximum extent of reaction. Reactions were for 60 s, except for mutant transcripts with guanosine (2 h). Different values for the same transcript and analogue correspond to different RNA preparations. Multiple determinations with the same RNA preparation were averaged. Splicing conditions were as in Table 1, and kinetic analysis as in Fig. 3. Synthesis of 2-aminopurine: thioguanosine (Sigma) was directly reduced to 2-aminopurine ribonucleoside with Raney nickel²⁰. The reaction was followed by ultraviolet spectroscopy: absorption at 340 nm decreased and that at 306 nm increased. A fine yellow precipitate was recovered and analysed. The ultraviolet spectrum was identical to that published by Fox *et al.*²⁰, and the NMR spectrum clearly indicated that the starting thioguanosine had been quantitatively (>99%) reduced. Synthesis of 2,6-diaminopurine riboside (an initial sample was a gift of L. Orgel): guanosine was quantitatively converted to 2,6-diaminopurine riboside in a silylation-amination procedure using hexamethyldisilazane and ammonia in the presence of a Lewis acid, trimethylsilyl-trifluoromethanesulphonate²¹. The product was recrystallized three times from boiling water to yield fine white crystals. Its NMR spectrum was identical to that found by Muraoka²² and showed no evidence of contamination by guanosine.

immediately 3' to a G residue^{17,18}. But our results show that there cannot be two completely separate binding sites, because the same mutation affects the G-residue requirement for both 5' and 3' splicing. A further argument in favour of a single G-binding site in Group I introns is that all cyclization reactions involving mismatched combinations of nucleotides and binding site are subject to competition by the better-matched free nucleoside. For instance, excised intron molecules carrying the A·U mutation circularize slowly, and this reaction is inhibited by 2AP but not by guanosine.

The interaction between guanosine and the highly conserved G·C base pair in the core of the ribozyme constitutes our first glimpse of how an RNA enzyme binds a mononucleotide substrate. Similar experiments with guanosine analogues and mutant introns should identify the bases involved in stacking interactions and the groups recognizing the ribose. □

Received 21 August; accepted 17 October 1989.

- Cech, T. R. *et al. Cell* **27**, 487-496 (1981).
- Cech, T. R. *Gene* **73**, 259-271 (1988).
- Price, J. V. & Cech, T. R. *Genes Dev.* **2**, 1439-1447 (1988).
- Tanner, N. K. & Cech, T. R. *Biochemistry* **26**, 3330-3340 (1987).
- Bass, B. L. & Cech, T. R. *Nature* **308**, 820-826 (1984).
- Bass, B. L. & Cech, T. R. *Biochemistry* **25**, 4473-4477 (1986).
- Davies, R. W. *et al. Nature* **300**, 719-724 (1982).
- Been, M. D. & Cech, T. R. *Nucleic Acids Res.* **13**, 8389-8408 (1985).
- Michel, F. & Dignon, B. *EMBO J.* **2**, 33-38 (1983).
- Waring, R. B. *et al. J. molec. Biol.* **167**, 595-605 (1983).
- Cech, T. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 3903-3907 (1983).
- Yarus, M. *Science* **240**, 1751-1758 (1988).
- Yarus, M. *Biochemistry* **28**, 980-988 (1989).
- Kay, P. S. & Inoue, T. *Nature* **327**, 343-346 (1987).

- Heinemann, O. & Saenger, W. *Nature* **299**, 27-31 (1982).
- Kay, P. S. *et al. EMBO J.* **7**, 3531-3537 (1988).
- Inoue, T. *et al. J. molec. Biol.* **189**, 143-165 (1986).
- Van der Horst, G. & Tabak, H. F. *EMBO J.* **6**, 2139-2144 (1987).
- Doudna, J. A. *et al. Cold Spring Harb. Symp. quant. Biol.* **52**, 173-180 (1987).
- Fox, J. J. *et al. J. Am. Chem. Soc.* **80**, 1669-1675 (1958).
- Vorbruggen, H. & Krolkiewicz, K. *Liebigs Ann. Chem.* **4**, 745-761 (1976).
- Muraoka, M. *Chem. pharmacol. Bull.* **29**, 3449-3454 (1981).
- Burke, J. M. *et al. Nucleic Acids Res.* **15**, 7217-7221 (1987).
- Szostak, J. W. *Nature* **322**, 83-86 (1986).
- Milligan, J. F. *et al. Nucleic Acids Res.* **15**, 8783-8798 (1987).

ACKNOWLEDGEMENTS. We thank Christian Schneider, Steven A. Benner and Nassim Usman for help with chemical syntheses, and Andrew Ellington for discussions. This work was supported by Hoechst AG, the NSF (R.G.) and the NIH (D.P.B.).

The barrier to recombination between *Escherichia coli* and *Salmonella typhimurium* is disrupted in mismatch-repair mutants

Christiane Rayssiguier*, David S. Thaler[†] & Miroslav Radman^{*‡}

* Institut Jacques Monod, 2 Place Jussieu, Tour 43, 75251 Paris-Cedex 05, France

[†] Biology Department, University of Utah, Salt Lake City, Utah 84112, USA

[‡] Laboratory of Molecular Genetics, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709, USA

The requirement for DNA sequence homology in generalized genetic recombination is greatly relaxed in bacterial *mutL*, *mutS* and *mutH* mutants deficient in mismatch repair. In such mutants, intergeneric recombination occurs efficiently between *Escherichia coli* and *Salmonella typhimurium*, which are ~20% divergent in DNA sequence. This finding has implications for speciation, for regulating recombination between diverged repeated sequences, and for hitherto difficult interspecies hybridizations.

Conjugation

Conjugation between *E. coli* Hfr donors and *S. typhimurium* F⁻ recipients deficient in DNA restriction (see Table 1 for strain description and the legend of Table 2 for the protocol) was conducted for 40 min on the surface of nutrient agar plates. Mating was interrupted and the bacteria were replated to select for *xyl*⁺ or *met*⁺ *S. typhimurium* recombinants under conditions that prevented the growth of the *E. coli* donor cells. When the recipient cells were deficient in methyl-directed mismatch repair, that is, *mutH*, *mutL*, *mutS* or *mutU*, recombinants were recovered much more frequently than when the recipients were *mut*⁺ (Table 2). The increase in the formation of recombinants is clearly *recA* dependent. (The relatively weak effect of *mutU* (leading to a lack of helicase II function)²³⁻²⁵ was not due to the complementation by the *E. coli* *mutU*⁺ gene—Table 2.) The co-inheritance of *xyl* and *met* markers, separated by ~13% of the bacterial genome, indicates that large fragments of the *Escherichia* genome can be integrated into the *Salmonella* recipient genome, although less frequently than in the homo-specific *E. coli* cross (Table 3).

Perhaps the simplest interpretation of these results is that intergeneric recombinants arise through gene replacement, the

GENERALIZED genetic recombination provides a means for the transfer or exchange of genetic information between homologous regions of DNA, as well as for the repair of DNA damage. The fidelity of recombination depends on the precise formation of a hybrid or heteroduplex joint consisting of two complementary DNA strands, one contributed by each of the recombining molecules¹⁻⁴. The hybrid joint is a key structure for aligning parental sequences in genetic recombination. Presumably because of the need to maintain chromosomal integrity, recombination *in vivo* requires nearly perfect homology and is abolished by as little as 10–20% sequence divergence^{5,6}. But heteroduplex DNA formation catalysed *in vitro* by the *E. coli* RecA protein tolerates up to 30% sequence mismatches⁷. Several lines of evidence indicate that the strand exchange reaction is edited by mismatch repair proteins supposed to abort or inactivate heteroduplex intermediates containing mismatched base pairs^{8,9}. Karimova *et al.*¹⁰ were first to propose that recombination between highly diverged sequences might be differentially inhibited by the mismatch repair system (owing to overlapping excision tracts¹¹⁻¹³); this idea has since been independently proposed by others^{8,9,14,15}. Studies on recombination and mismatch repair in *Streptococcus pneumoniae*^{16,17} and *E. coli*^{14,18} indicate that the marker effects on the efficiency of homologous recombination are often caused by the abortion of heteroduplex intermediates with one or few mismatches⁸. Thus, the mismatch repair system might prevent recombination between partially homologous sequences and provide a functional barrier to interspecies recombination^{8,9}. We have now tested the prediction that mutants defective in mismatch repair can become proficient in interspecies recombination by using members of two related bacterial genera, *E. coli* and *Salmonella typhimurium*, which are about 80% homologous in DNA sequence¹⁹. Mutations in the *mutL*, *mutS* or *mutH* mutator genes inactivate methyl-directed mismatch repair of errors in DNA replication²⁰⁻²²; we show that these mutations can increase intergeneric recombination in conjugational and transductional crosses by up to 1,000-fold. Hence, mismatch recognition and the mismatch repair system can act as a barrier to recombination between DNAs of different species.

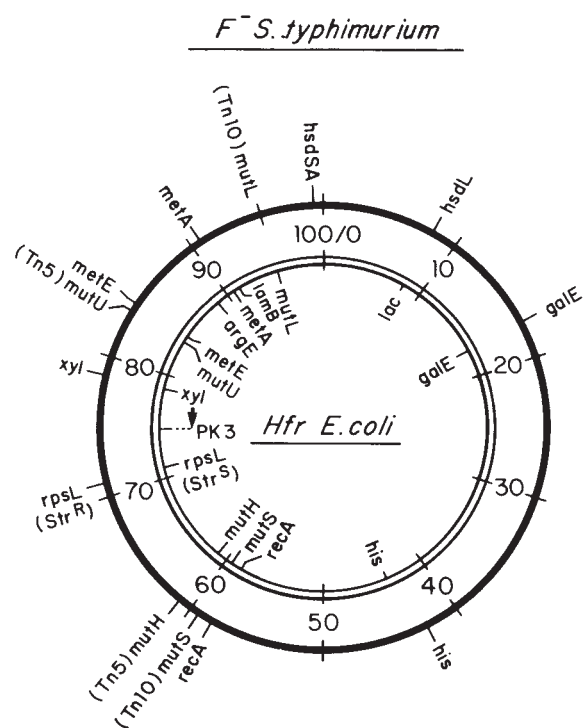


FIG. 1 Genetic maps of *E. coli* HfrPK3 and *S. typhimurium* LT2 F⁻ used in the experiments described here. The origin of transfer of the Hfr PK3 is indicated by an arrow (from refs 54–56).

TABLE 1 Genotypes and origins of bacterial strains

		Reference
<i>Escherichia coli</i>		
F ⁻ AB1157: <i>thr1 leu6 proA2 his4 thi1 argE3 lacY1 galK2 ara14 xyl5 mt1 tsx33 str31 supE44</i>		54
Hfr PK3: <i>thr1 leuB6 thi1 lacY1 azi15 tonA21 supE44</i>		55
Hfr PK3 <i>mutU</i> (<i>uvr D260::Tn5</i>)		This work*
<i>Salmonella typhimurium</i>		
F ⁻ SL4213: <i>galE496 metA22 metE551 rpsL120 xyl404</i> (Fels2) <i>H1b nml H2-enx ilv452 hsdSA29 hsdL6</i>		56
F ⁻ SL4213 <i>mutH::Tn5</i>		57
F ⁻ SL4213 <i>mutL::Tn10</i>		57
F ⁻ SL4213 <i>mutS::Tn10</i>		57
F ⁻ SL4213 <i>mutU::Tn5</i>		57
F ⁻ SL4213 <i>mutH::Tn5 recA1</i>		This work†
F ⁻ SL4213 <i>mutL::Tn10 recA1</i>		This work†
F ⁻ SL4213 <i>mutS::Tn10 recA1</i>		This work†

* Made by P1 transduction⁵⁸ from *E. coli* GW7303 (AB1157 *uvrD260::Tn5*), selection for kanamycin resistance (40 µg ml⁻¹) and screening for mutator and UV-sensitive phenotype.

† Made by P22 *int3* Hft transduction from *S. typhimurium* TT13479 (*recA1 srl203::Tn10d-Cam*—which confers chloramphenicol resistance)³⁴, selection for chloramphenicol resistance (10 µg ml⁻¹) and screening for sensitivity to ultraviolet.

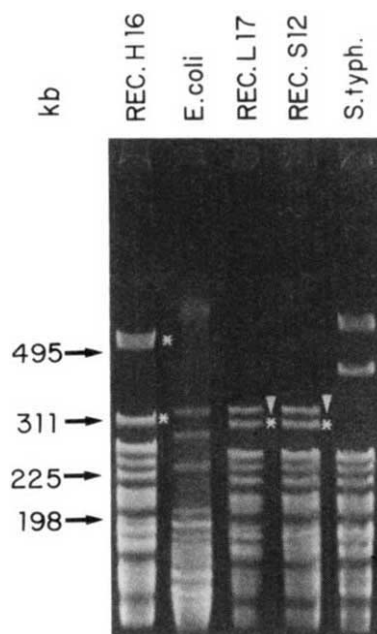
usual outcome of intraspecies mating. It is possible for phenotypic recombinants to arise through other mechanisms however, such as (1) *recA*-dependent circularization of the incoming DNA, perhaps using blocks of homology such as those found in ribosomal RNA *rrn* genes, and perpetuation of the circle as an autonomous plasmid (because the conjugated region includes the *E. coli oriC* origin of replication) and (2) formation of interspecies duplications involving recombination between sequences common to both organisms²⁶⁻²⁹.

Phenotypic recombinants resulting from plasmid or duplication formation should be less stable than recombinants resulting from gene replacement. To test for the stability of the *xyl*⁺ recombinants, colonies from each cross were streaked on indicator plates. Sectoring of the originally red colonies to white (the phenotype of the recipient before conjugation) or to dark red (the phenotype expected if multiple copies of the gene are present) indicates an unstable phenotypic recombinant. The fractions of stable recombinants are given in Table 3 and range

from 1 of 35 for *mutS* recipients to 19 of 35 for *mutU* recipients and 33 of 34 for *mut*⁺ recipients. If the recombinant frequencies in Table 2 are corrected for the fraction of unstable recombinants, the frequencies of stable recombinants from *mut*⁻ recipients range from 2×10^{-5} (*mutU*) to 7×10^{-4} (*mutL*), an increase of one to three orders of magnitude over the frequencies obtained with *mut*⁺ recipients. The highest frequencies of recombinants came from *mutL* and *mutS* recipients which could be a result of the involvement of the MutL and MutS proteins in the first step of the mismatch repair process³⁰. For example, it could be that the binding of these mismatch repair proteins to mismatch-containing DNA is sufficient to disrupt recombination even under conditions of complete DNA methylation. Alternatively, it could be that the incoming strands are undermethylated (perhaps newly synthesized)³¹ or that formation of partially homologous heteroduplex DNA involves extensive DNA synthesis in the flanking regions creating hemimethylated regions which would be substrates for normal mismatch repair activity.

FIG. 2 Restriction pattern of intergeneric recombinants from *E. coli* Hfr and *S. typhimurium* F⁻ crosses. Three stable *xyl*⁺ *met*⁺ recombinants (REC) were selected from crosses between *E. coli* Hfr PK3 and *S. typhimurium* F⁻ recipients which were either *mutH* (for REC.H16), *mutL* (for REC.L17) or *mutS* (for REC.S12). For REC.L17, the original *S. typhimurium mutL::Tn10* locus was replaced by the *mutL*⁺ locus from *E. coli*, as shown genetically in Table 3. Δ, Parental *S. typhimurium* F⁻ DNA fragments lost in recombinants; ▽, DNA fragment from recombinant genomes co-migrating with a parental *E. coli* Hfr donor DNA fragment; *, Non-parental DNA fragments. Fragment size markers are yeast *S. cerevisiae* no. 334 chromosomes IX, III, VI and I. Kb, kilobases.

METHODS. Bacteria (10 ml) were grown in Luria Bertani medium to $1.5-2 \times 10^8$ cells ml⁻¹, chloramphenicol was added to 180 µg ml⁻¹ and incubation was continued for 1 h. The bacteria were chilled, pelleted at 4 °C, resuspended in 10 ml solution I (10 mM Tris-HCl, pH 7.6, 1 M NaCl), pelleted again, and resuspended in 1.5 ml solution I. They were warmed to 40 °C and an equal volume of 1.5% low-melting-point agarose was added. The mixture was put in Tygon tubing, allowed to solidify, and extruded into an equal volume of lysis solution I (6 mM Tris-HCl, pH 7.6, 1 M NaCl, 100 mM EDTA, pH 7.5, 0.5% Brij 35, 0.2% deoxycholate, 0.5% sarcosyl, 1 mg ml⁻¹ lysozyme, 0.2 mg ml⁻¹ bovine pancreatic RNase) and incubated overnight at 37 °C with gentle shaking. (With these amounts, about 25 plugs with 0.5 µg DNA per plug can be made.) The plugs were then put in an equal volume of lysis solution II (0.5 M EDTA, pH 9.3, 1% lauroyl sarcosine, 1 mg ml⁻¹ proteinase K) and incubated for 2 days at 50 °C. The lysis solution II was replaced and the plugs were stored at 4 °C until used. Plugs were then put in 15 ml TAFE buffer (10 mM Tris buffer, 4.37 mM glacial acetic acid, 0.5 mM free-acid EDTA) at room temperature for 1-2 h with gentle inversion of the tubes. They were then put in new TAFE buffer with 5 mM phenylmethylsulphonyl fluoride (PMSF) and kept for another hour in the same conditions and then successively in 7 ml TE (10 mM Tris-HCl, pH 7.6, 1 mM EDTA, pH 8) for 15 min with gentle inversion of the tubes and in *SpeI*-restriction buffer containing 2 mM PMSF for 10 min at 37 °C. Ten units of *SpeI* enzyme (New England Biolabs) were then added to each plug, which was incubated overnight at 37 °C. The plugs were then put in transverse alternating-field electrophoresis



(TAFE) buffer and then in wells of 1.3% agarose gel and sealed with 1.5% agarose in TE. Electrophoresis in a Beckman Gene Line TAFE system was at 220 V with two switch pulses of 25 s each for 21 h at 14 °C. The gel was then stained in 0.5 µg ml⁻¹ ethidium bromide, destained, and illuminated with ultraviolet light. The low-melting-point agarose, the agarose and the TAFE buffer were Beckman Gene Line products. The experiment was done in the Beckman Instruments Spinco Division in Palo Alto.

TABLE 2 Conjugational crosses between *Escherichia coli* and *Salmonella typhimurium*

<i>E. coli</i> Hfr donor	<i>S. typhimurium</i> F ⁻ recipient	<i>E. coli</i> F ⁻ recipient	<i>xyl</i> ⁺ per Hfr	<i>metE</i> ⁺ <i>metA</i> ⁺ per Hfr
<i>mut</i> ⁺	<i>mut</i> ⁺		9 × 10 ⁻⁷	2 × 10 ⁻⁶
<i>mut</i> ⁺	<i>mutS</i>		2 × 10 ⁻³	2 × 10 ⁻³
<i>mut</i> ⁺	<i>mutL</i>		3 × 10 ⁻³	2 × 10 ⁻³
<i>mut</i> ⁺	<i>mutH</i>		6 × 10 ⁻⁵	2 × 10 ⁻⁴
<i>mut</i> ⁺	<i>mutU</i>		5 × 10 ⁻⁵	9 × 10 ⁻⁶
<i>mut</i> ⁺	<i>mut</i> ⁺ <i>recA</i>		7 × 10 ⁻⁸	
<i>mut</i> ⁺	<i>mutS</i> <i>recA</i>		< 2 × 10 ⁻⁸	
<i>mut</i> ⁺	<i>mutL</i> <i>recA</i>		3 × 10 ⁻⁷	
<i>mut</i> ⁺	<i>mutH</i> <i>recA</i>		4 × 10 ⁻⁸	
<i>mutU</i>	<i>mutU</i>		3 × 10 ⁻⁵	
<i>mutU</i>	<i>mutL</i>		3 × 10 ⁻³	
<i>mut</i> ⁺		<i>mut</i> ⁺	3 × 10 ⁻¹	
<i>mutU</i>		<i>mut</i> ⁺	2 × 10 ⁻¹	

Heterospecific crosses involved *E. coli* Hfr PK3 or PK3 *mutU* and *S. typhimurium* F⁻ SL4213 *mut*⁺, *mutH*, *mutL*, *mutS*, *mutU* or F⁻ SL4213 (*mut*⁺ or *mut*) *recA1* (see Table 1). Homospecific crosses involved *E. coli* Hfr PK3 or PK3 *mutU* and *E. coli* F⁻ AB1157. Log-phase bacteria were mixed in a 1:1 (Hfr:F⁻) ratio and immediately deposited on a sterile 0.45-μm pore size filter (Schleicher and Schuell) which was incubated on prewarmed, rich-medium agar. After 40 min at 37 °C, the conjugants were resuspended in 10⁻² M MgSO₄ and viciously separated by vortexing. The exconjugants were then plated on M63 synthetic medium⁵⁸ either lacking methionine to select for *met*⁺ recombinants or containing xylose as the sole carbon source to select for *xyl*⁺ recombinants. The Hfr donor cells were doubly counterselected by 25 μg ml⁻¹ streptomycin and, in the case of intergeneric conjugation, by withholding threonine and leucine. Amino acids were added to a final concentration of 100 μg ml⁻¹, thiamine to 30 μg ml⁻¹ and xylose to 0.4%. Because of late appearance of intergeneric recombinants from *mut*⁺ and *mutU* crosses, all recombinants were scored after 88 h of incubation. Recombination frequencies (*xyl*⁺ or *metE*⁺*metA*⁺) are expressed per Hfr donor after subtracting unmated revertants. Each number represents the mean of two to four independent experiments (the range was up to fourfold).

This conclusion follows from the evidence that the MutH protein recognizes and cleaves unmethylated GATC sequences in the course of methyl-directed mismatch repair^{30,32}. Note, however, that the requirement for unmethylated GATC sequences and MutH protein is alleviated when strand discontinuities flank the mismatch-containing region^{30,32}.

To assess whether *Escherichia coli* *mut*⁺ genes replaced *Salmonella* *mut*⁻ genes, we monitored the transposon (Tn) insertions in the *Salmonella* *mut* genes: if the functional *E. coli* *mut*⁺ gene replaced the resident Tn-bearing *mut* gene, then the recombinant will have lost both its Tn-borne drug resistance and its mutator character. This was so for seven out of eight stable *xyl*⁺ *met*⁺ recombinants from the cross with *S. typhimurium* *mutL*::Tn10 as the recipient, indicating that except in one recombinant, *Escherichia* sequences had replaced *Salmonella* sequences for ~5 min beyond the last selected gene (*metA*) into the *mutL* region (see Fig. 1). None of 27 unstable *xyl*⁺ *met*⁺ recombinants from the same cross lost drug resistance or the mutator phenotype (Table 3).

In the cross with *S. typhimurium* *mutU*::Tn5 as the recipient, the *mutU* gene was replaced in all of the 19 stable *xyl*⁺ recombinants (17 of which were also *met*⁺: the *xyl* and *met* loci flank *mutU*, see Fig. 1). Also, out of 16 unstable recombinants, 3 also showed *mutU* replacement, indicating that coincident replacement and duplication occurred. None of the *xyl*⁺ *met*⁺ recombinants from *mutS*::Tn10 and *mutH*::Tn5 crosses showed transposon loss or *mut*-gene replacement, because mating was interrupted at 40 min and the *mutS*-*mutH* chromosomal region of genes was therefore unable to be transferred (see Fig. 1). The stability of *xyl*⁺ recombinants from *mut*⁺ and *mutU* crosses does not prove that gene replacement occurs because heterospecific duplications could be stable owing to very rare homeologous recombination in *mut*⁺ and *mutU* strains (Table 2).

Additional evidence for the physical linkage of *E. coli* and *S. typhimurium* DNA sequences comes from the study of restriction enzyme digests of parental and recombinant genomes. DNA of the parental strains and three stable recombinants (one each from *mutH*, *mutL* and *mutS* crosses) was cut with *SpeI* enzyme (Fig. 2). The two largest fragments (together, ~1 megabase) from parental *S. typhimurium* recipient are lost in all three recombinants and each of the recombinants has at least one nonparental fragment, which presumably contains an interspecies junction. Two of the recombinants have a large fragment which co-migrates with a large *E. coli* fragment.

Transduction

Transduction experiments were conducted by using phage P22 to move either *E. coli* or *S. typhimurium* histidine operon sequences between two *S. typhimurium* strains. The various recipients carried insertion sequences—IS200 (ref. 33) for the experiments of Table 4A and Tn10-*dCam* (ref. 34) for those in Table 4B—in the *hisD* gene and also a deletion in an unlinked locus (*proA*, *B*) essential for proline biosynthesis. The donor for heterospecific transduction had a deletion of most of the *his* locus³⁵, but was a histidine prototroph because it harboured a 150-kilobase (kb) *his*⁺ F' episome from *E. coli*. Thus, the *his* locus in the donor strain was uniquely encoded by *E. coli* sequences, whereas the *pro* locus was the normal *S. typhimurium* sequence. Transducing lysate for homospecific crosses was prepared on wild-type *S. typhimurium* and thus contained *S. typhimurium* sequences at both the *his* and the *pro* locus. A comparison of the *hisD*⁺:*proA*⁺ *B*⁺ transduction ratios from the two lysates gives a measure of the efficiency of interspecies recombination relative to intraspecies recombination. The results (Table 4) are qualitatively similar to the results of conjugation experiments, in that *mutL* and *mutS* have the strongest effect and *mutU* the smallest effect. Because the *E. coli* sequences used in these transduction experiments were obtained from an episome present in *S. typhimurium*, the effect of the mismatch repair system must be related only to sequence differences between the species rather than to a species-specific modification of the DNA.

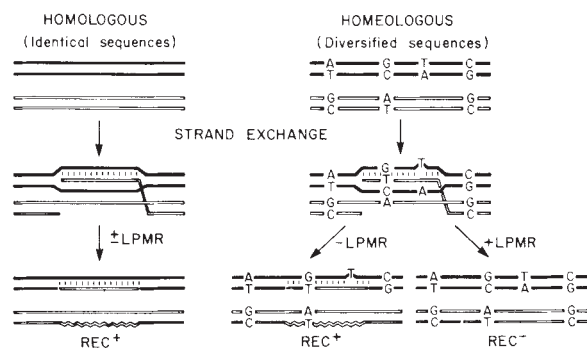


FIG. 3 Molecular model for selective anti-recombination by LPMR (long-patch mismatch repair) proteins. Homologous (identical) sequences recombine with the same efficiency in the presence or absence of active LPMR proteins (Table 4, and ref. 9). Homeologous (diversified) sequences start to recombine by forming a heteroduplex containing mismatches caused by sequence differences in the two parental DNAs. If LPMR proteins are not functional, homeologous recombination proceeds to form recombinant products (REC⁺ phenotype). If LPMR proteins are functional mismatches in the heteroduplex regions are recognized, the reaction is reversed by the removal or destruction of the heteroduplex intermediate, and no viable recombinants arise (REC⁻ phenotype). Although irrelevant to the mechanism of anti-recombination, the initiation of strand exchange is arbitrarily shown in a single-strand invasion mode⁶³. Only base-pair differences between the parental molecules are indicated: the hybrid region in this sketch (shown with dashes indicating hydrogen bonding) contains one G-T mismatch plus one unpaired T residue (as in a frameshift mutation).

Evidence for heterospecific gene replacement is shown in Table 4B. The *hisD* allele in the recipient contains a Tn10 derivative that confers chloramphenicol resistance. Only 1 of 318 *hisD*⁺ transductants retained chloramphenicol resistance. All of 400 *proA*⁺*B*⁺ transductants retained chloramphenicol resistance. A *Mud-lacZ* fusion to *hisD* is likewise replaced during heterospecific recombination by P22 transduction (data not shown). Because the homospecific recombination by P22 transduction occurs by integration of double-stranded DNA³⁶, the *mutH* effect in our experiments could have been due to repair synthesis generating new undermethylated DNA in *cis* to the incoming DNA.

Discussion

The results reported here indicate that the suppression of mismatch repair can allow interspecies recombination between sequences that have as much as 20% sequence divergence. Evidence for an inhibitory action of the *E. coli* mismatch repair system on recombination between diverged sequences has been found in three other experimental systems. First, by using a very short patch (vsp) mismatch repair system as a means to measure the formation of heteroduplex DNA intermediates *in vivo* it was found that the *mutH-mutU* system suppresses recombinant formation in bacteriophage λ crosses owing to the presence of two genetic markers¹⁴. Second, carrying out crosses between a phage- λ vector and a plasmid, each bearing a mouse VH immunoglobulin gene sequence, in a *mutS* host was found to increase recombination between partially homologous, but not between identical, VH genes¹⁵. Finally, Karimova *et al.* presented evidence indicating that recombination between two highly diverged lambdoid prophages can be enhanced in mutator hosts¹⁰.

We favour proposals in which recombination between partially homologous sequences involves a reversible heteroduplex intermediate that is detected and aborted through a process that involves mismatch recognition^{8,9} (see Fig. 3). This is the stage at which genetic recombination is limited in transformation of *S. pneumoniae* by the Hex mismatch repair system^{16,17}. The alternative proposal involves the mismatch-stimulated killing¹¹⁻¹³ of the products of recombination between diverged sequences^{10,15}. Although the mechanism by which mismatch repair proteins act to suppress recombination has not been elucidated, it seems that they edit not only the products of DNA replication²⁰⁻²², but also intermediates in the process of genetic recombination.

There is evidence that mismatch repair also operates in fungal³⁷, amphibian³⁸ and mammalian³⁹ cells. The average sequence divergence among individual copies of the major dispersed repetitive sequence families in mammalian genomes, such as the *Alu* sequences⁴⁰, is similar to the sequence divergence between the genomes of *E. coli* and *S. typhimurium*. The mismatch repair system could act to prevent inter- and intrachromosomal rearrangements by suppressing recombination between misaligned dispersed repetitive sequences and thus play an important part in maintaining gene order in bacteria⁴¹ and in eukaryotes. Similarly, if the recombination suppressing effect of diverged sequences extends into neighbouring regions of perfect homology, then the high sequence polymorphism among homologous chromosomes in humans⁴² could contribute to the suppression of mitotic recombination without compromising the ability of the cells to use recombinational DNA repair involving exchanges between sister chromatids of identical sequences (sister chromatid exchange). Mitotic recombination leads to the expression of deleterious recessive mutations which are often associated with pathological states such as cancer⁴³.

Some regulation of the mismatch repair system, or of the extent of heteroduplex DNA formation, would presumably be necessary during meiosis to allow pairing and recombination between homologues in spite of their sequence polymorphism. If meiosis becomes deficient in chromosome pairing, for

example, when sequence divergence is sufficiently extensive and a mismatch repair system is sufficiently active, chromosome nondisjunction will result in sterility⁴⁴. Fertility can then be restored through inbreeding or selfing, thus avoiding sequence divergence, or by transiently inactivating the mismatch repair system^{45,46}. The result would be the emergence of a new species. Thus, mismatch repair could be a molecular mechanism for maintaining a genetic barrier between species that have diverged, but not drastically, on the DNA sequence level, and could accomplish the genetic separation (that is, reproductive isolation) required for evolution of species in the absence of geographical isolation⁴⁷. The acquisition of dispersed noncoding DNA, such as repetitive sequences and introns, could accelerate speciation by allowing an increased rate of genomic sequence divergence and subsequent meiotic and mitotic recombinational isolation of genes, chromosomes and species. An alternative chromosomal mechanism for speciation is the accumulation of large chromosomal rearrangements. An extensive study of a marsh rat population showed that chromosomal rearrangements, however, have a smaller role in reproductive isolation than that previously believed⁴⁸.

Mismatch repair apparently helps to prevent all classes of genetic alterations; it prevents point mutations by correcting replication errors^{16,20-22} and chromosomal rearrangements by preventing recombination between sequences of only limited homology (limited in size⁴¹ or extent—this work and ref. 15). Modulation of the mismatch repair system^{45,46} provides an effective way to control the rate of 'vertical' evolution (occurring by means of new point mutations and intrachromosomal rearrangements) and 'horizontal' evolution (occurring through recombination with diverged sequences).

This work shows that suppression of mismatch repair allows interspecies recombination between bacteria that have been diverging for ~150 million years¹⁹. Hybrids were formed between *E. coli* and *S. typhimurium* that are in fact new species with only partial fertility in crosses with each of two parental

TABLE 3 Analysis of *xyl*⁺ and *met*⁺ recombinants from *E. coli* Hfr × *S. typhimurium* F[−] conjugational crosses

F [−] recipient	Co-inheritance of markers		<i>xyl</i> ⁺ recombinants	
	% <i>met</i> ⁺ among <i>xyl</i> ⁺	% <i>xyl</i> ⁺ among <i>met</i> ⁺	Stable: unstable	Transposon loss
<i>S. typhimurium</i> :				
<i>mut</i> ⁺	7.7 (389)	5.8 (412)	33:1	
<i>mutL::Tn10</i>	3.6 (412)	9.4 (412)	8:27	7/8, 0/27
<i>mutS::Tn10</i>	8.7 (412)	6.6 (375)	1:34	0/1, 0/34
<i>mutH::Tn5</i>	10.7 (412)	1.9 (412)	7:28	0/7, 0/28
<i>mutU::Tn5</i>	2.5 (412)	22.0 (412)	19:16	19/19, 3/16
<i>E. coli</i> :				
<i>mut</i> ⁺	63.1 (412)*	18.9 (412)*	20:0	

To determine co-inheritance of markers, *xyl*⁺ or *met*⁺ recombinants were replicated on media selecting for *met*⁺ or *xyl*⁺ bacteria, respectively, and scored after 2 days of incubation. Total numbers of colonies analysed are given in parentheses. The stability of *xyl*⁺ recombinants was tested by growing each colony in a rich liquid medium and streaking on 1% xylose McConekey agar, on which *xyl*⁺ bacteria give red colonies and *xyl*[−] bacteria white colonies. For each recombinant, a red colony was grown and streaked a second time. The recombinants that then gave only red colonies were scored as stable, whereas those giving some white, dark red or sectorised colonies were scored as unstable. Numbers are of recombinants scored. Transposon loss was tested by replicating each of the recombinants previously scored for stability onto a rich medium containing tetracycline (12.5 µg ml^{−1}) for Tn10 or kanamycin (50 µg ml^{−1}) for Tn5, and scoring resistant colonies after one day of incubation. The first and second fractions represent the number of stable and unstable recombinants, respectively, expressing drug sensitivity. All recombinants that lost the original Tn10 or Tn5 were shown in subsequent tests to have also lost the mutator phenotype (data not shown).

* An *arg*⁺ marker used instead of *met*⁺; the two are closely linked.

TABLE 4 Transduction crosses

A. Mutator alleles enhance the ability of *E. coli* sequences to donate histidinol prototrophy to *S. typhimurium*

Recipient strain mut allele	Heterospecific crosses <i>hisD</i> ⁺ : <i>E. coli</i> <i>proA</i> ⁺ <i>B</i> ⁺ : <i>S. typhimurium</i>	Homospecific crosses <i>hisD</i> ⁺ : <i>S. typhimurium</i> <i>proA</i> ⁺ <i>B</i> ⁺ : <i>S. typhimurium</i>
	<i>hisD</i> ⁺ / <i>proA</i> ⁺ <i>B</i> ⁺ ($\times 10^4$)	<i>hisD</i> ⁺ / <i>proA</i> ⁺ <i>B</i> ⁺
<i>mut</i> ⁺	4	320
<i>mutL</i> no. 1	355	180
<i>mutL</i> no. 2	147	150
<i>mutS</i> no. 1	696	130
<i>mutS</i> no. 2	451	160
<i>mutU</i> no. 1	20	460
<i>mutU</i> no. 2	10	520
<i>mutH</i> no. 1	37	180
<i>mutH</i> no. 2	170	190

B. Interspecific transduction by allelic replacement

Recipient strain mut allele	Plate contains:	Plate selects for:	Mean no. of colonies
<i>mut</i> ⁺	Pro, Hol	His ⁺	5
	Pro, Hol, Cam	His ⁺ , Cam ^r	0
	His, Arg	Pro ⁺	478
	His, Arg, Cam	Pro ⁺ , Cam ^r	498
<i>mutS</i>	Pro, Hol	His ⁺	168
	Pro, Hol, Cam	His ⁺ , Cam ^r	0
	His, Arg	Pro ⁺	1,344
	His, Arg, Cam	Pro ⁺ , Cam ^r	1,110

In part A, the recipient for P22 transduction was *hisD984::IS200 proAB47* where *hisD::IS200* is a nonreverting transposable element insertion³³ and *proAB47* is a nonreverting deletion of genes essential for proline prototrophy⁵⁹. The donor P22 Hft *int3* lysate⁶⁰ carrying *E. coli hisD* sequences and *S. typhimurium proAB* sequences was grown on strain TR5360 from John Roth's strain collection; although strain TR5360 contains *his646*, a deletion covering the *hisD* gene³⁵, it is a histidinol prototroph because it also carries *E. coli* F' 150 which includes the entire *E. coli his* operon⁶¹. All *S. typhimurium* strains are in the LT2 background, although the *proAB47* allele was first isolated in LT7 and subsequently moved into LT2. The P22 donor lysate with *S. typhimurium* sequences at both the *his* and the *pro* loci was prepared on LT2. No. 1 and 2 refer to two independent *mut(H, L, S* or *U)* derivatives of the initial recipient strain; these were obtained by P22-mediated transduction (as in Table 1) of mutator alleles from SL4213 *mut* strains (Table 1). Selection was conducted by spreading 0.1 ml of donor lysate and 0.1 ml of a fresh overnight culture of recipient cells on a minimal-glucose plate; the plates were scored after 2–3 days at 37 °C. Aliquots of the same overnight culture and P22 donor lysate were used for both the histidinol (His⁺) and the proline (Pro⁺) selections. When His⁺ was selected, the plates contained proline and histidinol; for Pro⁺ selection, the plates contained both histidine and arginine; the arginine prevents *argD* mutants from acting as indirect suppressors of proline auxotrophy⁶². In part B, the recipient was *proAB47 hisD::Tn10d-Cam* (which confers chloramphenicol resistance as well as histidinol auxotrophy³⁴). *E. coli* sequences donated *hisD*⁺ and *S. typhimurium* sequences donated *proAB*⁺. Transductions were conducted as for part A. Undiluted donor lysate (0.1 ml) was used for transductions to histidinol prototrophy and 0.1 ml of a 10-fold dilution of the lysate was used for transductions to proline prototrophy. The numbers in the last column are the means of independent experiments; ranges were less than twofold. Precise and nearly precise excision of transposons are increased in mutator backgrounds⁴¹ but the insertion alleles used here are not susceptible to mutator-enhanced excision generating prototrophs: neither the *hisD984::IS200* nor the *Tn10d-Cam* allele reverted, even in mutator backgrounds (data not shown). Cam, chloramphenicol; Cam^r, chloramphenicol-resistant phenotype; Hol, histidinol.

nonpathogens (such as *E. coli*) could allow the production of hybrid species suitable for live vaccines. Interspecies transfer of interesting alleles should facilitate genetic analysis of related species. Easy selection for gene replacements, duplications and amplifications after conjugation could provide useful microorganisms in which the duplicated heterospecific region is stabilized by converting recombinants to *mut*⁻ (Table 3). In addition, interspecies genetics could facilitate studies on the structure and function of hybrid operons (Table 4), genes and proteins⁵ and the production of proteins with altered activity, stability or specificity. For example, recombination between diverged genes in *mutL*, *S* bacteria should create complex heteroduplex regions which are substrates for diverse specialized vsp mismatch repair systems (see ref. 9 for a review) generating complex patchwork systems of the two parental sequences⁴⁹. By recombining and patchworking functionally equivalent but sequence-diverged genes cloned from different species, it should be possible to create new diversified genes that still conserve their basic motif and yet alter stability or specificity of the gene products. This strategy for protein 'engineering' could be similar to hyperconversion mechanisms generating diversity in the expressed genes by their recombination with diverged silent (pseudo)genes (see ref. 50 for a review), as in the case of immunoglobulin genes⁵¹ and genes coding for variable surface proteins in parasites^{52,53}. □

Received 13 August; accepted 12 October 1989.

- Holliday, R. *Genet. Res.* **5**, 282–304 (1964).
- Whitehouse, H. L. K. *Nature* **199**, 1034–1040 (1963).
- Fox, M. S. & Allen, M. *Proc. natn. Acad. Sci. U.S.A.* **52**, 412–419 (1964).
- Meselson, M. in *Heritage from Mendel* (ed. Brink, R. A.) 3–16 (University of Wisconsin Press, 1965).
- Schneider, W. P., Nichols, B. P. & Yanofsky, C. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2169–2173 (1981).
- Shen, P. & Huang, H. V. *Genetics* **112**, 441–457 (1986).
- Das Gupta, C. & Radding, C. M. *Nature* **295**, 71–73 (1982).
- Radman, M. in *Genetic Recombination* (eds Kutcherlapati, R. & Smith, G. R.) 169–192 (American Society for Microbiology, Washington DC, 1988).
- Radman, M. *Genome* **31**, 68–73 (1989).
- Karimova, G. A., Grigoriev, P. S. & Rybchin, V. N. *Molekularnaia Genetika Mikrobiologia i Virusologia* **10**, 29–34 (1985).
- Glickman, B. W. & Radman, M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1063–1067 (1980).
- Wagner, R. et al. *Cold Spring Harb. Symp. quant. Biol.* **49**, 611–615 (1984).
- Doutriaux, M.-P., Wagner, R. & Radman, M. *Proc. natn. Acad. Sci. U.S.A.* **83**, 2576–2578 (1986).
- Jones, M., Wagner, R. & Radman, M. *Cell* **50**, 621–626 (1987).
- Shen, P. & Huang, H. V. *Molec. Gen. Genet.* **218**, 358–360 (1989).
- Claverys, J. P. & Lacks, S. A. *Microbiol. Rev.* **50**, 133–165 (1986).
- Mejean, V. & Claverys, J. P. *Molec. Gen. Genet.* **197**, 467–471 (1984).
- Feinstein, S. I. & Low, K. B. *Genetics* **113**, 13–33 (1986).
- Ochman, H. & Wilson, A. C. in *Escherichia coli and Salmonella typhimurium* (ed. Neidhardt, F. C.) 1649–1654 (American Society for Microbiology, Washington DC, 1987).
- Radman, M. & Wagner, R. A. *Rev. Genet.* **20**, 523–538 (1986).
- Modrich, P. A. *Rev. Biochem.* **56**, 435–466 (1987).
- Meselson, M. in *The Recombination of Genetic Material* (ed. Low, B. K.) 91–113 (Academic, New York, 1988).
- Kushner, S. R. et al. *UCLA Symp. molec. cell. Biol.* **11**, 153–159 (1983).
- Hickson, I. D. et al. *Molec. Gen. Genet.* **190**, 265–270 (1983).
- Kumura, K. & Sekiguchi, M. *J. biol. Chem.* **259**, 1560–1565 (1984).
- Baron, L. S. et al. *Bacteriol. Rev.* **32**, 362–369 (1968).
- Middleton, R. B. & Mojica-A. T. *Adv. Genet.* **16**, 53–79 (1971).
- Anderson, P. R. & Roth, J. R. A. *Rev. Microbiol.* **31**, 473–505 (1977).
- Lehner, A. F. & Hill, C. W. *Genetics* **110**, 365–380 (1985).
- Lahue, R. S., Au, K. G. & Modrich, P. *Science* **245**, 160–164 (1989).
- Siddiqui, O. & Fox, M. S. *J. molec. Biol.* **77**, 101–123 (1973).
- Laengle-Rouault, F., Maenhaut-Michel, G. & Radman, M. *EMBO J.* **6**, 1121–1127 (1987).
- Lam, S. & Roth, J. R. *Cell* **34**, 951–960 (1983).
- Elliott, T. & Roth, J. R. *Molec. Gen. Genet.* **213**, 332–338 (1988).
- Hartman, P. E., Hartman, Z. & Stahl, R. C. *Adv. Genet.* **16**, 1–34 (1971).
- Ebel-Tsipis, J., Fox, M. S. & Botstein, D. *J. molec. Biol.* **71**, 449–469 (1972).
- Bishop, D. K., Andersen, J. & Kolodner, R. D. *Proc. natn. Acad. Sci. U.S.A.* **86**, 3713–3717 (1989).
- Brooks, P. et al. *Proc. natn. Acad. Sci. U.S.A.* **86**, 4425–4429 (1989).
- Brown, T. C. & Jiricny, J. *Cell* **54**, 705–711 (1988).
- Britten, R. J. et al. *Proc. natn. Acad. Sci. U.S.A.* **85**, 4770–4774 (1988).
- Lundblad, V. & Kleckner, N. *Genetics* **109**, 3–19 (1985).
- White, R. & Lalouel, J.-M. *Sci. Am.* **258**, 40–48 (1988).
- Weinberg, R. A. *Cancer Res.* **49**, 3713–3721 (1989).
- Murray, A. W. & Szostak, J. W. A. *Rev. cell. Biol.* **1**, 289–315 (1985).
- Damagnez, V., Doutriaux, M.-P. & Radman, M. *J. Bact.* **171**, 4494–4497 (1989).
- Schaaper, R. M. & Radman, M. *EMBO J.* **8**, 3511–3516 (1989).
- Barton, N. H., Jones, J. S. & Mallet, J. *Nature* **336**, 13–14 (1988).
- Nachman, M. W. & Myers, P. *Proc. natn. Acad. Sci. U.S.A.* **86**, 6666–6670 (1989).
- Abastado, J. P. et al. *Proc. natn. Acad. Sci. U.S.A.* **81**, 5792–5796 (1984).
- Maizels, N. *Trends Genet.* **5**, 4–8 (1989).
- Radman, M. in *Cellular Responses to DNA Damage* (eds Friedberg, E. C. & Bridges, B. A.) 287–298 (Liss, New York, 1983).
- Thon, G., Baltz, T. & Eisen, H. *Genes Dev.* **3**, 1247–1254 (1989).
- Seifert, H. S. et al. *Nature* **336**, 392–395 (1988).
- Bachmann, B. J. in *Escherichia coli and Salmonella typhimurium* (ed. Neidhardt, F. C.) 807–844 (American Society for Microbiology, Washington DC, 1987).

organisms. The utility of technologies allowing interspecies recombination is wide ranging. For example, attenuating the pathogenicity of *Salmonella*, *Shigella* or other pathogenic species by substituting appropriate genomic sequences from related

55. Kahn, P. L. *J. Bact.* **96**, 205-214 (1968).
 56. Sanderson, K. E. & Stocker, B. A. D. in *Escherichia coli and Salmonella typhimurium* (ed. Neidhardt, F. C.) 1220-1224 (American Society for Microbiology, Washington DC, 1987).
 57. Shanabrook, W. G., Belau, I. & Walker, G. C. *J. Bacteriol.* **147**, 827-835 (1981).
 58. Miller, J. H. *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, New York, 1972).
 59. Csoska, L. N. *Molec. Gen. Genet.* **182**, 82-86 (1981).
 60. Schmiegler, H. *Molec. Gen. Genet.* **119**, 75-88 (1972).
 61. Low, K. B. *Bact. Rev.* **36**, 587-607 (1973).
 62. Leisinger, T. in *Escherichia coli and Salmonella typhimurium* (ed. Neidhardt, F. C.) 345-351 (American Society for Microbiology, Washington DC, 1987).

63. Meselson, M. & Radding, C. M. *Proc. natn. Acad. Sci. U.S.A.* **72**, 358-361 (1975).

ACKNOWLEDGEMENTS. We thank D. Petranovic, R. D'Ari, C. Langley, M. Mergeay, D. Hillyard and J. W. Drake for advice and discussions, J. W. Drake, R. Wagner, and S. M. Rosenberg for advice on the manuscript, H. Huang and P. Shen for sharing unpublished results, and G.S. Walker, K. E. Sanderson (*Salmonella* Genetic Stock Center), J. R. Roth and J.-M. Louarn for bacterial strains. N. Avdalovic provided advice and hospitality to C. R. in the Beckman Instruments Spinco Division (Palo Alto). D.S.T. was supported by l'Association pour la Recherche sur le Cancer (ARC) during a visit to Paris, and in Salt Lake City by the Helen Hay Whitney Foundation. Work in Paris was supported by the CNRS, INSERM and ARC and in Salt Lake City by NIH grants to John Roth.

LETTERS TO NATURE

Extragalactic radiation and the ultra-high-energy cosmic-ray spectrum

F. W. Stecker

Laboratory for High Energy Astrophysics,
 NASA Goddard Space Flight Center, Greenbelt, Maryland 20771, USA

The high-frequency (submillimetre) side of the 'black-body' radiation spectrum has now become available for study from balloons and the soon-to-be-launched Cosmic Background Explorer (COBE) satellite. There is recent evidence of an unexpectedly large excess of radiation beyond the peak of the black-body curve¹, which seems to be generally isotropic² and therefore of extragalactic extent. Such radiation will have a significant effect on the propagation of ultra-high-energy cosmic rays. Because of the importance of this new discovery and the timeliness of the subject, I re-examine here the effect of extragalactic microwave and submillimetre-radiation fields on the ultra-high-energy cosmic-ray spectrum. Although the most detailed work^{3,4} so far is based on complex numerical simulations of individual interactions, I find that the general characteristics of the spectrum can be derived from fairly simple analytical arguments. I show that one can explore, by simpler and more general means, the various spectral features obtained by numerical calculations ('pile-up' bumps and 'dips'^{3,4}). I illustrate this technique with a newly derived lifetime-energy relation, based on the new submillimetre observations.

Semi-analytical methods can be useful in examining the implications of forthcoming data on background radiation fields

without computer simulations involving specific parameterizations. Such features can be obtained directly from a consideration of the average lifetime curves⁵⁻⁷ and the phase-space continuity equation. For simplicity, I consider that the highest-energy cosmic rays are protons. (Similar considerations can be extended to nuclei, but that case is complicated by a time-variable charge-dependent pair-production loss rate owing to the photodisintegration process⁸.) The effective lifetime of such cosmic rays against energy loss by photomeson production (or equivalent mean-free-path) on black-body radiation of temperature, T_B , is given by⁵

$$\tau^{-1}(E) = (1/4\pi)\gamma^2 h^3 c^2 \int_{\epsilon'_{th}/2\gamma}^{\infty} d\epsilon' (\exp(\epsilon'/kT_B) - 1) \times \int_{\epsilon'_{th}}^{2\gamma\epsilon'} d\epsilon' \sigma_{\gamma p}(\epsilon') K(\epsilon') \quad (1)$$

where γ is the proton Lorentz factor, ϵ' is the photon energy in the rest frame of the ultra-high-energy proton and the inelasticity, $K(\epsilon')$, is the mean energy lost by photomeson production per interaction.

Equation (1) for a pure black-body spectrum can easily be extended to include the recently discovered submillimetre-radiation field by using the parameterization of a pure black-body field and a further submillimetre gray-body radiation field⁸ of temperature $T_G = 4.9$ K and dilution factor 0.04. The cross-section data originally used in ref. 5 have been updated using recent compilations of the Particle Data Group⁹. Pair-production losses⁵ have also been included to obtain the total lifetime curve. Figure 1 shows the function $\tau(E)/E$, which is the inverse energy loss rate in GeV^{-1}s . The solid curve includes the effect of the submillimetre radiation and the dashed curve indicates the effect of the black-body radiation field only.

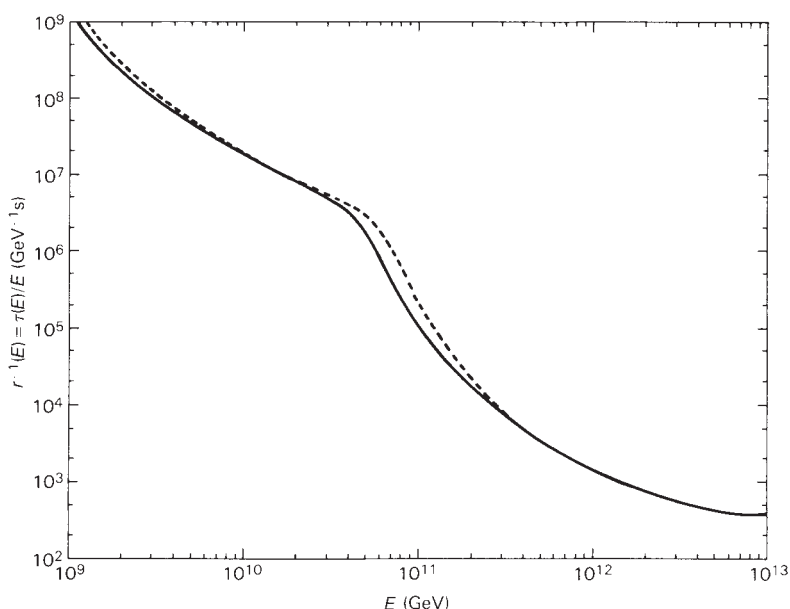


FIG. 1 The inverse energy loss rate $r^{-1}(E)$ produced by the black-body radiation field only (dashed line) and the black-body radiation field plus the submillimetre radiation field (solid line).

I note that the mean lifetime curve itself (which can be obtained from Fig. 1) is the source of much information. First, it indicates that if the sources of ultra-high-energy cosmic rays are of cosmological origin, from quasars or primordial galaxies for example, protons of energies larger than $\sim 10^{10}$ GeV will be degraded by energy losses from pair production and photomeson production^{10,11}. Second, protons in the energy range $\geq 10^{11}$ GeV should have a 'local' origin, most likely from sources within the local Virgo Supercluster, if the spectrum shows no attenuation effects at the highest energies⁵. The local-supercluster-origin model first suggested in ref. 5 has been developed elsewhere^{12,13}.

Figure 1 can be used as a starting point for a further examination of the spectral features to be expected, given an ultra-high-energy cosmic-ray source function $q(E, t)$. The quantity $\tau(E)$ defined in equation (1) should not be misinterpreted. It represents a characteristic energy loss time for the protons, not a decay time for catastrophic loss. In fact, because of the inclusion of the integration over the inelasticity term $K(\varepsilon')$, $\tau(E)$ can be rewritten so that a simple formalism that automatically conserves proton number can be used. To do this, I first define a function describing the mean energy loss rate from pair production and photomeson production interactions, $r(E) \equiv dE/dt$ so that

$$r(E) = E/\tau(E) \quad (2)$$

and the quantity τ/E plotted in Fig. 1 is actually the inverse of the mean proton energy loss rate. Conservation of proton number is then assured by the relation

$$\frac{\partial n(E, t)}{\partial t} + \frac{\partial}{\partial E} [n(E, t)r(E)] = q(E, t) \quad (3)$$

which describes a divergenceless flow of cosmic-ray protons in energy space in the case for which I assume spatial isotropy and no catastrophic losses. In this case, the cosmic-ray flux is directly proportional to the proton number density

$$I(E, t) = n(E, t)c/4\pi \quad (4)$$

Equations (3) and (4) have the solution¹⁴

$$I(E, t) = \frac{\tau(E)c}{4\pi E} \int_E^\infty dE' q\left(E', t - \int_E^{E'} \frac{dE''}{E''} \tau(E'')\right) \quad (5)$$

In equation (5), the quantity $\tau(E)/E$ has been substituted for the inverse of the absolute value of the energy loss rate, as follows from equation (2). Equation (5), which was obtained using Green's function techniques, clearly indicates that the cosmic-ray spectrum at time t is obtained from two factors. These are a term that is the inverse of the energy loss rate at energy E , which accounts for the 'trickle out' to energies below

E and is the factor given in Fig. 1, and a 'trickle down' input factor given by the integral over source spectra from earlier times

$$t' = t - \int_E^{E'} \frac{dE''}{E''} \tau(E'') \quad (6)$$

and higher energies $E' > E$, for which the initial energies have been degraded by energy losses. For example, it follows from the results in Fig. 1 that it takes $\sim 10^9$ yr for a $\sim 10^{11}$ GeV proton to be degraded to an energy of $\sim 4 \times 10^{10}$ GeV. It is the effect of the 'trickle down' integral that can cause a 'pile-up' in the spectrum.

The treatment I use here breaks down at energies $\leq 10^9$ GeV, the point at which the energy loss time becomes comparable to or greater than the Hubble time. For cosmic rays of truly cosmological origin, the age of which is a significant fraction of the age of the Universe, energy losses and dilution effects from the expansion of the Universe have a critical role. In this case, the first term in equation (3) must be replaced by a derivative with respect to redshift^{7,14}. I will not pursue the subject of the cosmological cosmic rays here as it is beyond the scope of this paper and does not seem to bear on the origin of the highest-energy cosmic rays⁵.

As illustrations of how this formalism can be used, I consider a steady-state solution of equation (3), and a 'burst' delta-function source spectrum that has been 'aged' by a specific amount (refs 3, 4).

To find the steady-state time-independent solution, I set the $\partial n/\partial t$ term in equation (3) equal to zero. Equation (5) then becomes

$$I(E) = \frac{\tau(E)}{E} \int_E^\infty dE' q(E') \quad (7)$$

which can be integrated if $q(E)$ is a power-law spectrum $KE^{-\Gamma}$. In this case

$$I(E) = \frac{q(E)\tau(E)}{\Gamma - 1} \quad (8)$$

Thus, the factor $\tau(E)$ has the role of a modulation factor on the initial source spectrum. It is as if the cosmic rays of energy E were built up over a time given by $\tau(E)/(\Gamma - 1)$. The flux at energy E becomes proportional to the energy loss time at energy E as well as the source flux at energy E .

Figure 2 shows the form of the spectrum obtained by an application of equation (8) in the case of an E^{-2} source spectrum such as might be expected if the highest-energy cosmic rays are produced by shock acceleration in extragalactic sources¹⁵. The spectrum is superimposed on a compilation of recent air-shower data¹⁶⁻²⁰. A 'pile-up' effect between 2 and 5×10^{10} GeV can

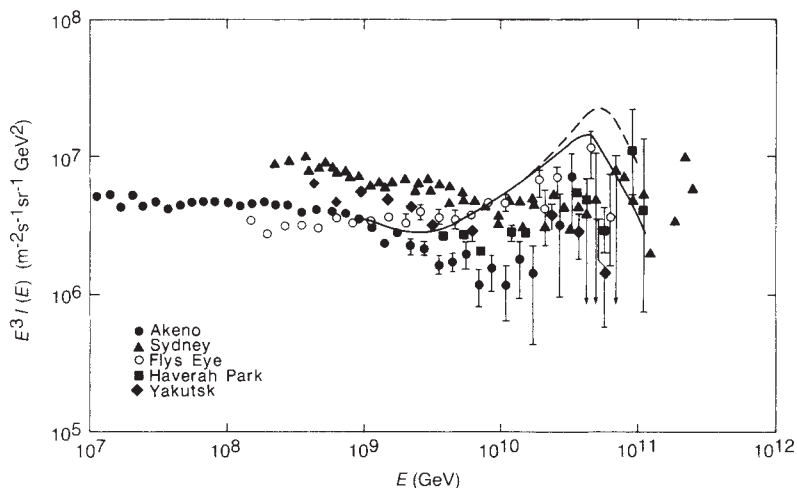


FIG. 2 The cosmic-ray flux spectrum derived from various air-shower results shown with the shape of the steady-state equilibrium spectrum obtained from equation (8). (Solid and dashed lines as in Fig. 1.)

clearly be seen in the theoretical curve, in qualitative agreement with earlier numerical results^{3,4}, although evidence of such an effect does not seem to be very clearcut in the collected data in this energy range, as can be seen from the figure. Indeed, interpretation of the highest-energy air-shower data has been hotly debated by the experimental groups involved¹⁷.

Of course, there is no *a priori* reason to favour a steady-state source mode. I therefore also consider the example of a 'burst' approximation as the other extreme. In my second example, the solution to equation (3) with $q(E, t) = q(E)\delta(t - t_0)$ is given by fixing $t' = t_0$ in equation (6). With this constraint, it follows from considering variations of the endpoints of the integral on the right-hand side that

$$\frac{\delta E}{\delta E'} = \frac{\tau(E')/E'}{\tau(E)/E} = \frac{r(E)}{r(E')} \quad (9)$$

so that if $r(E)/r(E') \ll 1$, protons originating in the energy range $\Delta E'$ will end up after time t_0 confined to the much smaller energy range ΔE and 'pile-up'. In that case, $t_0(E, E') \rightarrow t_0(E)$.

It follows that one can then define a 'pile-up energy' given by the inverse function $E_{pu}(t_0)$. As an example, I consider the simple case in which, in the energy range $4 \times 10^{10} \text{ GeV} \geq E \geq 10^{11} \text{ GeV}$, I can approximate $\tau(E)$ by a simple power-law $\tau(E) = \kappa E^{-\phi}$ where $\kappa = 2.3 \times 10^{39}$ with E in GeV and τ in years and $\phi \approx 2.8$. Then, for $t_0 = 3 \times 10^8 \text{ yr}$, I find an expected pile-up at

$$E_{pu} = [t_0 \phi / \kappa]^{-(1/\phi)} \approx 7 \times 10^{10} \text{ GeV} \quad (10)$$

in excellent agreement with the numerical results.

The formalism presented here can be used as a tool for further interpretation of the effects of submillimetre and far-infrared extragalactic background fluxes on ultra-high-energy source spectra, as new data on these background radiation fluxes become available. The results obtained using this technique should be of adequate accuracy, considering the magnitude of the various uncertainties in the air-shower data. This technique can thus complement numerical simulations made with specific assumptions. $\square$

Received 24 July; accepted 3 October 1989.

1. Matsumoto, T. *et al. Astrophys. J.* **329**, 567–571 (1988).
2. Bernstein, G. M. *et al. Astrophys. J.* **337**, L1–L4 (1989).
3. Puget, J. L., Stecker, F. W. & Bredekamp, J. H. *Astrophys. J.* **295**, 638–654 (1976).
4. Hill, C. T. & Schramm, D. N. *Phys. Rev. D* **31**, 564–580 (1985).
5. Stecker, F. W. *Phys. Rev. Lett.* **21**, 1016–1018 (1968).
6. Blumenthal, G. R. *Phys. Rev. D* **1**, 1596–1602 (1970).
7. Berezhinsky, V. S., Grigor'eva, S. I. & Zatesepin, G. T. *Astrophys. Space Sci.* **36**, 17–30 (1975).
8. Daly, R. A. *Astrophys. J.* **333**, L1–L3 (1988).
9. Particle Data Group *Physics Lett.* **B204**, 1–486 (1988).

10. Greisen, K. *Phys. Rev. Lett.* **16**, 748–750 (1966).
11. Zatesepin, G. T. & Kuz'min, V. A. *JETP Lett.* **4**, 78–80 (1966).
12. Berezhinsky, V. S. & Grigor'eva, S. I. in *Proc. 16th Int. Cosmic Ray Conf. (Kyoto)* **2**, 81–85 (1979).
13. Giler, M., Wdowczyk, J. & Wolfendale, A. W. *J. Phys. G* **6**, 1561–1573 (1980).
14. Stecker, F. W. *Cosmic Gamma Rays* (Mono Book, Baltimore, 1971).
15. Sorrell, W. H. *Astrophys. J.* **323**, 647–656 (1987).
16. Baltraitis, R. M. *et al. Phys. Rev. Lett.* **54**, 1875–1877 (1985).
17. Linsley, J. in *Proc. 19th Int. Cosmic Ray Conf. (La Jolla)* **9**, 475–486 (1985).
18. Kristiansen, G. B. in *Proc. 19th Int. Cosmic Ray Conf. (La Jolla)* **9**, 487–497 (1985).
19. Winn, M. M. *et al. in Proc. 19th Int. Cosmic Ray Conf. (La Jolla)* **9**, 499–507 (1985).
20. Teshima, M. *et al. in Proc. 20th Int. Cosmic Ray Conf. (Moscow)* **1**, 397–400 (1987).

Helium abundance and asymmetry in the wind from the precursor to supernova 1987A

D. A. Allen*, W. P. S. Meikle† & J. Spyromilio*†

* Anglo-Australian Observatory, PO Box 296, Epping, New South Wales 2121, Australia

† Blackett Laboratory, Imperial College of Science, Technology and Medicine, Prince Consort Road, London SW7 2BZ, UK

To explain why the progenitor of supernova 1987A exploded as a blue supergiant star, models of stellar evolution require that there must have been extensive mixing into the hydrogen-rich envelope of materials from deeper, helium-rich regions^{1–3}. Estimates of the helium abundance in the outer layers, which can be obtained from infrared spectroscopy of the stellar wind from the progenitor, can help to determine the extent of the mixing. The initial ultraviolet flash from the surface of supernova 1987A ionized the circumstellar gas to produce narrow emission lines, which were first seen at ultraviolet wavelengths⁴ where the supernova's photospheric emission quickly faded. Later, optical lines were also reported⁵. The ultraviolet lines indicated electron densities of $\sim 10^4 \text{ cm}^{-3}$, far greater than expected for the wind shed by the precursor star, Sk – 69 202, in either its blue supergiant or preceding red supergiant phase⁶. However, the interaction of the two winds could produce a shell of the requisite density^{7,8}. From observations of the infrared helium triplet at 1,083 nm we infer that the red supergiant wind was strongly asymmetric and that part of it had a high helium abundance.

In the wind interaction^{7,8}, masses as high as a few solar masses can be compressed to densities $> 10^4 \text{ cm}^{-3}$ as a shell of radius about one light yr ($\sim 10^{18} \text{ cm}$) and thickness 10^{15} cm . But its morphology, which depends on the importance of instabilities⁶ and/or interaction with the interstellar medium, might exhibit a significant departure from spherical symmetry. The initial ultraviolet flash is predicted to have released 10^{56} – 10^{57} photons

(refs 9–11), sufficient to ionize almost 1 solar mass of gas. These figures were confirmed by the ultraviolet line ratios during the first 400 days (refs 6, 11). The evolution of the flux and angular scale of the emission from the ionized gas is determined by the effects of morphology, recombination time and light travel time⁷. At densities of 10^4 – 10^5 cm^{-3} , recombination timescales are months or years. The ultraviolet line emission can be understood in terms of a spherically symmetric ionized interaction shell⁶. However, we find that the helium in the shell does not have a spherical distribution.

Asymmetry in the optical emission from supernova 1987A has already been found. In 1987, speckle observations^{12,13} briefly revealed a companion (the 'mystery spot') ~ 20 light days from the supernova at position angle $\sim 195^\circ$. At much the same position angle, spectropolarimetry¹⁴ indicated an elongation or asymmetry of the supernova. When the ejecta themselves became resolvable¹⁵ they too were extended along position angle 20 – 200° . More recently, velocity gradients were seen in the circumstellar gas¹⁶; the greatest recession velocity is at $\sim 190^\circ$. Complex arc-shaped structures extending to 3 arcsec from the supernova have also been reported¹⁷.

We have undertaken a near-infrared spectroscopic study of supernova 1987A using the Anglo-Australian Telescope¹⁸. Since May 1988, we have recorded a narrow component to the 1,083-nm He I line. Measurements were obtained on six occasions between days 434 and 737. A square aperture of side 3.5 arcsec was used, always centred on supernova 1987A. Figure 1 shows two of the spectra. During this time, the narrow line showed little evidence of variation, remaining at an intensity of $(15 \pm 2) \times 10^{-13} \text{ erg cm}^{-2} \text{ s}^{-1}$. But, as noted below, the source of much of this emission was offset from the supernova, and lay near the aperture edge. Consequently the measured intensity may have been affected by the pointing and tracking stability.

On day 737 we explored the spatial distribution of the narrow He I emission feature¹⁹. We set the grating so that one of the 16 detectors in the spectrometer (FIGS) received radiation from the narrow component while several straddled the broad component. By repeated raster scanning of the telescope we thus created simultaneous maps of the two components (Fig. 2). The

seeing was ~ 1.5 arcsec in diameter. Although the broad component, which is due to the supernova itself, originated in an unresolved point source, approximately half of the total intensity of the narrow component arose in a compact source or 'knot' displayed by 0.8 ± 0.3 arcsec at position angle $200 \pm 15^\circ$. The knot diameter was ≤ 1.5 arcsec (1.2 light yr). A faint northerly extension to the emission was also seen. We also made north-south line scans through supernova 1987A, using the same spectral and spatial sampling, and a similar displacement in declination was seen (Fig. 3). Clearly, most of the emission in the narrow component had a circumstellar origin.

Several observations indicate that the compact source lay behind the supernova, and so had been observable for less than the supernova lifetime. First, circumstellar gas at the source position angle is receding⁸. Second, contemporary mid-infrared measurements (P. F. Roche, D. K. Aitken and C. H. Smith, personal communication) revealed supernova-illuminated dust lying in a compact region displaced from the supernova and, to within the errors, coincident with the He I knot. Because dust cools virtually instantaneously, its emission maps out the present location of energy from the light-curve peak. Considerations of geometry, light travel time and temperature indicate that the dust must lie behind the supernova^{20,21}. Third, direct optical continuum imaging around day 750 indicates scattering from dust with a region of enhanced surface brightness at about the same sky position as the He I knot¹⁷. Geometry and light-travel-time effects again point to a location behind the supernova. Optical speckle observations²² in a continuum bandpass also show a source at about the same position from days 644 to 697, although there seems to be some disagreement between direct and speckle imaging as to the position and magnitude of the brightest region^{17,23,24}. Fourth, the high density derived below shows that the source did not comprise the material emitting the ultraviolet lines during the first 1.1 years.

We have constructed a simple model of the compact He I source consisting of a slowly moving cloud of hydrogen and helium only. We assume that it was fully ionized by the ultraviolet flash, and that light travel time across it was unimportant. Our model predicts almost constant He I 1,083-nm emission per unit volume from 3 months to 1 year after initial ionization.

From the constancy of the observed intensity we conclude that the mass of emitting gas did not increase significantly during the period of observation, so that by day 737, we had been viewing the same cloud for at least 0.8 yr. If we adopt a viewing time of 1 yr since ionization, the cloud lies 0.8 light yr from the supernova.

We have allowed for recombination, collisional effects and ionization of neutral atoms by photons released in the recombination of He²⁺. We take the He I line to be optically thin and use atomic parameters from refs 25–30. An electron temperature of $\sim 25,000$ K was inferred from approximately contemporary intensity ratios of the narrow 436.3, 495.9 and 500.7-nm [O III] lines (R. Stathakis, personal communication). A higher temperature was found initially³¹, but cooling would have been rapid¹¹. Only some of the [O III] emission had the same sky position as the knot¹⁷, and so we do not know if the [O III]-derived temperature is appropriate for the helium cloud. We therefore adopted 30,000 K as a conservative upper limit for the cloud temperature—a lower temperature would increase the inferred helium overabundance. Our only free parameters were the electron density N_e and the helium mass fraction Y . We corrected for interstellar reddening^{32,33} and included contributions of He II to the Balmer lines.

We have compared the model with our He I observations and with the intensities of narrow H α , H β and He II 469-nm lines obtained from A. P. S. Crotts (personal communication) and from the Anglo-Australian Observatory supernova archive (R. Stathakis, personal communication). Around day 750, the total narrow H α intensity from the region within 3 arcsec of the supernova was 9×10^{-13} erg cm⁻² s⁻¹. The total narrow H β intensity was 0.3 of this. The He II 469-nm intensity decreased by about a factor of four between days 613 and 756 and had approximately one-third of the total H β intensity on day 756. These intensities are un-reddened.

From images obtained around day 750, it is estimated (A. P. S. Crotts, personal communication) that 0.3 ± 0.1 of the narrow H α flux came from within the 1.5-arcsec maximum extent of the He I knot. In comparing the model and observations, we therefore made the conservative assumption that 0.3 of the hydrogen emission originated from the cloud. For $T = 30,000$ K,

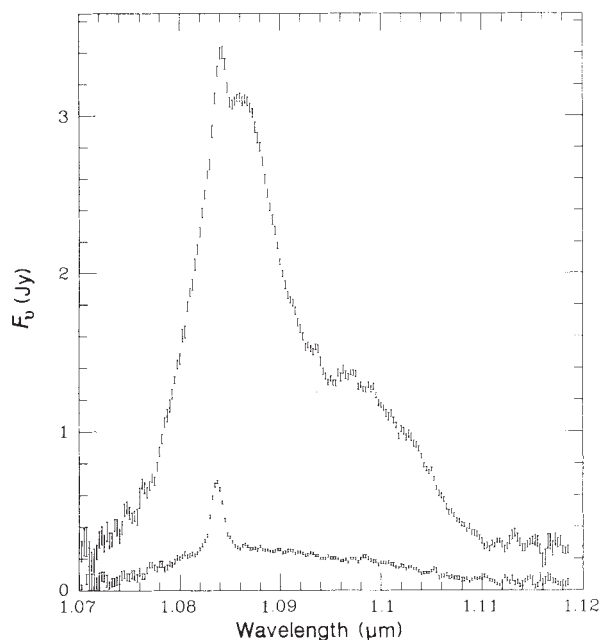


FIG. 1 Spectra of supernova 1987A taken in September 1988 and January 1989. The broad component is helium and hydrogen emission from the supernova ejecta. The narrow component is helium in the circumstellar shell.

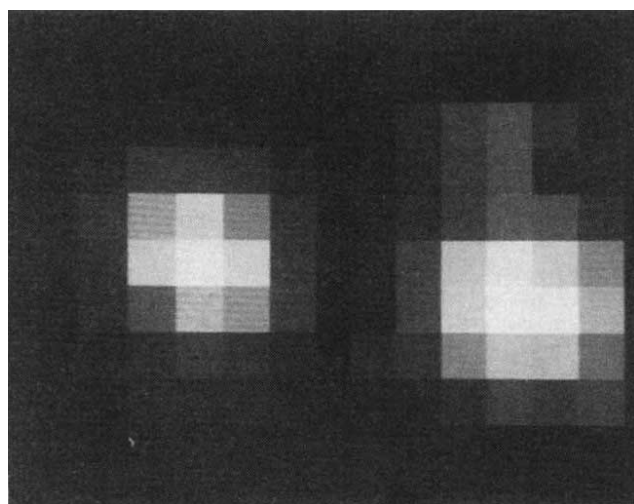


FIG. 2 Grey-scale representations of maps of the broad (left) and narrow (right) components seen in Fig. 1. North is at the top and east to the left. The focal plane was sampled on a square grid, with 0.7-arcsec pixels, through a 1.4-arcsec square aperture. The contribution of the broad component was removed from the observed flux in the narrow component by scaling from a fully sampled spectrum. The resulting image shows the distribution of the narrow component alone. Both maps have been automatically scaled to cover the full tonal range.

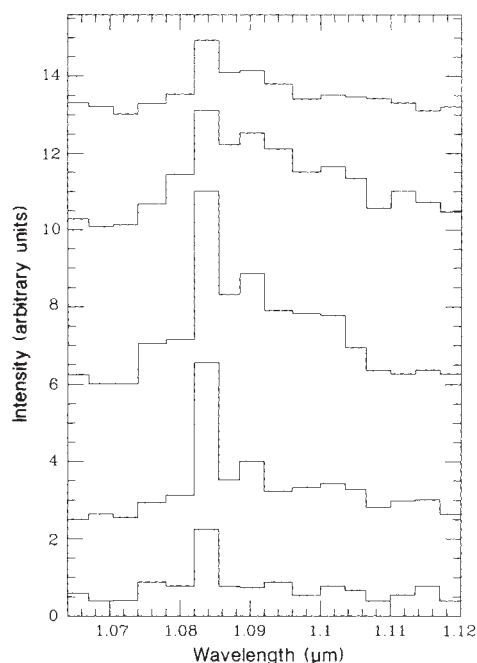


FIG. 3 Six undersampled spectra from ~ 200 co-added declination scans through supernova 1987A, taken at 0.7-arcsec spacing. The narrow component peaks in the second and third spectra from the bottom (south), and the broad component in the third and fourth.

the model satisfactorily reproduces the observed fluxes if $N_e = 10^5 \text{ cm}^{-3}$ and Y is in the range 0.4–0.5. The corresponding $N(\text{H})$ is also $\sim 10^5 \text{ cm}^{-3}$. The 1,083-nm He I line arises both directly by recombination of ionized helium and from collisional excitation of the metastable 2^3S level. At $N_e \sim 10^5$, the latter dominates and the line is strengthened relative to the recombination case by more than an order of magnitude²⁵. The cloud mass density is about 10 times larger than the typical value for the whole shell obtained from ultraviolet observations⁶. From the He I line intensity we derive a total mass of ionized gas of $\sim 0.01 M_\odot$ in the cloud. This is a factor of 5–10 less than the shell mass inferred from the ultraviolet lines^{6,11}.

In the plane of the sky the shell extended to ~ 2 light yr from the supernova^{16,17,31} and had lower density⁶. The cloud could not have been produced by the second (blue supergiant) wind, for if it had expanded into an isotropic shell it could only have accumulated high densities at larger radii. The compact He I source may be an interstellar cloud that successfully resisted the momentum flux of the red-supergiant wind. However, it is unlikely that an interstellar cloud with so high a helium abundance would lie within 1 light yr of the supernova. An alternative is a bipolar red supergiant wind¹⁶, possibly also interacting with the interstellar medium. A symmetrical bipolar form is incompatible with our data because the frontal lobe would now be seen at age 2 yr, yielding a 1,083-nm intensity 60% that of the rear lobe. Moreover, any such frontal lobe must have been less dense than the rear if the density is to be compatible with the ultraviolet line ratios. We therefore interpret the cloud as part of an asymmetric red-supergiant outflow. An interacting outflow behind the supernova may account for the extended mid-infrared and continuum optical emission^{17,20}. Crotts *et al.*¹⁷ suggest that the circumstellar shell is near a dense flat sheet, possibly of interstellar origin, lying behind the supernova and point out that the He I knot lies close to where the shell and sheet intersect.

Extensive mass loss and mixing of the progenitor envelope have been invoked to account for its blue-red-blue evolution before explosion and to explain the high N/C and N/O ratios inferred from the ultraviolet data^{1–3}. With the further constraint

that the progenitor envelope was massive, Saio *et al.*² find that complete mixing of the hydrogen-rich envelope together with dredge-up of about one solar mass of material from the helium core is required, giving a surface helium abundance ~ 0.45 . Our observations support this model.

Finally we note that the cloud subtends up to 2 sr at the supernova so the total number of photons capable of doubly ionizing helium and radiated isotropically in the ultraviolet flash must have been at least 10^{55} , in agreement with model predictions and ultraviolet observations^{9–11}. □

Received 12 June; accepted 3 October 1989.

- Maeder, A. *ESO Workshop on the Supernova 1987A, Conf. Workshop Proc.* **26**, 251–269 (1987).
- Saio, H., Nomoto, K. & Kato, M. *Nature* **334**, 508–510 (1988).
- Woosley, S. E., Pinto, P. A. & Weaver, T. A. *Proc. astr. Soc. Aust.* **7**, 355–370 (1988).
- Wamsteker, W., Gilmozzi, R. & Cassatella, A. *Int. astr. Un. Circ.* **4410** (1987).
- Wampler, J. *Int. astr. Un. Circ.* **4541** (1988).
- Fransson, C. *et al. Astrophys. J.* **336**, 429–441 (1989).
- Chevalier, R. A. *Nature* **332**, 514–516 (1988).
- Chevalier, R. A. *ESO Workshop on the Supernova 1987A, Conf. Workshop Proc.* **26**, 481–494 (1987).
- Dopita, M. A., Meapheringham, S. J., Nulsen, P. & Wood, P. R. *Astrophys. J.* **322**, L85–L89 (1987).
- Lundqvist, P. & Fransson, C. *ESO Workshop on the Supernova 1987A, Conf. Workshop Proc.* **26**, 495–502 (1987).
- Fransson, C. & Lundqvist, P. *Astrophys. J.* **341**, L59–L62 (1989).
- Nissen, P., Papaliolios, C., Karovska, M. & Noyes, R. *Astrophys. J.* **320**, L15–L18 (1987).
- Meikle, W. P. S., Matcher, S. J. & Morgan, B. L. *Nature* **329**, 608–611 (1987).
- Cropper, M. S. *et al. Mon. Not. R. astr. Soc.* **231**, 695–722 (1988).
- Papaliolios, C. *et al. Nature* **338**, 565–566 (1989).
- Wood, P. R. & Faulkner, D. J. *Int. astr. Un. Circ.* **4739** (1989).
- Crotts, A. P. S., Kunkel, W. E. & McCarthy, P. J. *Astrophys. J.* (in the press).
- Meikle, W. P. S., Allen, D. A., Spyromilio, J. & Varani, G. F. *Mon. Not. R. astr. Soc.* **238**, 193–223 (1989).
- Allen, D., Meikle, P. & Spyromilio, J. *Int. astr. Un. Circ.* **4747** (1989).
- Roche, P. F., Aitken, D. K., Smith, C. H. & James, D. D. *Nature* **337**, 533–535 (1989).
- Felten, J. E. & Dwek, E. *Nature* **339**, 123–125 (1989).
- Karovska, M., Nissen, P., Papaliolios, C., Standley, C. & Heathcote, S. *Int. astr. Un. Circ.* **4749** & **4752** (1989).
- Crotts, A. & Kunkel, W. E. *Int. astr. Un. Circ.* **4753** (1989).
- Heathcote, S., Suntzeff, N. & Walker, A. *Int. astr. Un. Circ.* **4753** (1989).
- Osterbrock, D. E. *Astrophysics of gaseous nebulae and active galactic nuclei* (University Science Books, Mill Valley, 1989).
- Jacobs, V. L. & Burke, P. G. *J. Phys.* **B5**, 2272–2281 (1972).
- Bell, K. L., Kingston, A. E. & Taylor, I. R. *J. Phys.* **B6**, 2271–2279 (1973).
- Hata, J. & Grant, I. P. *J. Phys.* **B14**, 2111–2124 (1981).
- Clegg, R. E. S. *Mon. Not. R. astr. Soc.* **229**, 31P–39P (1987).
- Clegg, R. E. S. & Harrington, J. P. *Mon. Not. R. astr. Soc.* **239**, 869–883 (1989).
- Wampler, E. J. & Richichi, A. *The Messenger* **52**, 14–16 (1988).
- Dopita, M. A. *Proc. astr. Soc. Aust.* **7**, 344–351 (1988).
- West, R. M., Lauberts, A., Jørgensen, H. E. & Schuster, H.-E. *Astr. Astrophys.* **177**, L1–L2 (1987).

ACKNOWLEDGEMENTS. We thank R. Clegg, A. Crotts, R. Learner, R. Stathakis and P. Storey for discussion. We are also grateful to A. Crotts and R. Stathakis for providing data before publication.

Photoisomerization of OCIO: a possible mechanism for polar ozone depletion

Veronica Vaida*, Susan Solomon†, Erik C. Richard*, Eckart Rühl*‡ & Anne Jefferson*

* Department of Chemistry, University of Colorado, Boulder, Colorado 80309, USA

† Aeronomy Laboratory, NOAA, Boulder, Colorado 80303, USA

‡ Permanent address: Institut für Physikalische und Theoretische Chemie, Freie Universität Berlin, Takustrasse 3, D-1000 Berlin 33, FRG

CONNECTIONS between polar ozone depletion and halocarbon chemistry have been established by a number of studies^{1–10}. Recent attempts to account quantitatively for the observed rate of ozone decline in Antarctica in terms of known photochemical processes have not been entirely successful, and it seems that further chemical ozone-depleting mechanisms may be needed, particularly if the transport of ozone into the polar regions competes with chemical losses. Spectroscopic and photochemical data indicate that photolysis of OCIO may provide a further ozone loss mechanism that has not previously been considered. Here we report laboratory studies of OCIO spectroscopy and photoproducts which suggest

that atomic Cl and O₂ are formed to some extent in the photodissociation process. This evidence points towards possible photoisomerization to the unstable species ClOO, (or at least to a similar metastable intermediate) probably by way of the ²B₂ excited state of OCIO, thus reinforcing the idea that photolysis of OCIO may contribute to polar ozone depletion.

The discovery by Farman *et al.*¹ in 1985 that a precipitous drop in Antarctic ozone had occurred has been the impetus for intensive studies of the Antarctic stratosphere. Those studies quickly established that the Antarctic ozone decline occurs largely in September²⁻⁴ when the total column of Antarctic ozone drops by ~50% as compared with historical values obtained in the 1960s. A number of workers suggested that the decline in ozone might be related to increasing chlorofluorocarbon abundances¹, and noted that the unique depletion in the Antarctic spring could result from the combination of extremely cold temperatures which produce polar stratospheric clouds (PSCs), and the presence of sunlight to drive photochemical processes⁵⁻⁷. Detailed studies of the stratospheric composition during September (in 1987, in particular) revealed that the chemistry of the Antarctic stratosphere had been greatly perturbed, leading to large abundances of chlorine radicals that can destroy ozone quite readily in the sunlit atmosphere^{8,9}. These studies provided firm links between the ozone loss and halocarbon chemistry.

Although it is clear that the highly unusual chlorine chemistry induced by heterogeneous reactions occurring on the surfaces of polar stratospheric clouds can account for a substantial fraction of the ozone loss, it is difficult to be more quantitative because of uncertainties in both the laboratory and field measurements. A number of catalytic cycles involving chlorine monoxide radicals participate in the chemical destruction of ozone in the sunlit Antarctic stratosphere. Recent studies⁸⁻¹¹ are in general agreement that formation and photolysis of the ClO dimer is the dominant mechanism for Antarctic ozone destruction, at least during 1987. Coupling between chlorine and bromine represents a second important ozone loss pathway, the rate limiting step of which is



Although reactions involving the ClO dimer are likely to dominate ozone loss in Antarctica under cold conditions and when ClO abundances are high, the bromine-chlorine mechanism assumes greater importance at warmer temperatures or when ClO enhancements are modest. It is therefore likely to play a major and possibly dominant role in Arctic ozone loss, but also contributes ~20-30% of the Antarctic ozone depletion according to present understanding. Barrett *et al.*⁸ suggested that the formation and photolysis of the ClO dimer could explain virtually all of the observed Antarctic ozone loss in 1987 based on their measurements of chlorine monoxide, but recent measurements for the reaction rate constant of ClO with itself to form the dimer (Cl₂O₂) have reduced the apparent importance of this mechanism. There is now considerable question as to whether or not currently understood chemical mechanisms can explain

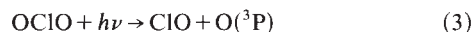
the observed rate of Antarctic ozone loss⁹⁻¹¹. One study⁹ estimated that only 65% of the total observed Antarctic ozone loss rate could be accounted for through the photochemistry suggested to date, whereas others conclude that the photochemical loss may be sufficient to account for the entire ozone decline.

Many studies have implicitly or explicitly assumed that the Antarctic vortex acts entirely as a 'containment vessel' in which chemistry takes place in total isolation. Any downward or poleward transport of ozone into the depleted region will probably require considerable increases in the rate of chemical destruction to account fully for the observed ozone decrease^{12,13}.

The interest in polar ozone has led to a fuller understanding of many of the chemical mechanisms that occur in the lowest part of the stratosphere. In spite of this rapid progress, it seems likely that other unidentified chemical processes could contribute to the ozone decline. In particular, it is widely accepted that a second product channel of the reaction between ClO and BrO leads to formation of OCIO

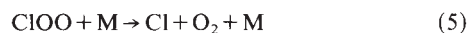


This process is not expected to account for a net ozone loss, as rapid gas-phase photolysis of OCIO has hitherto been assumed to produce exclusively atomic oxygen, which will reform ozone by the reaction with molecular oxygen.



Note the contrast between the chemistry of OCIO, which is thermally stable, and its isomer, ClOO, which is extremely unstable.

If photolysis of OCIO were to involve the following two-step process



then formation of OCIO would be a chemical pathway to ozone destruction. If all OCIO photolysis were to proceed through photoisomerization, then the ozone loss rate that is due to the coupling of chlorine and bromine would roughly double (as the rate constants for reactions (1) and (2) are believed to be comparable at cold temperatures). Given the extremely labile nature of ClOO, it is expected to be very short lived and, probably cannot be observed directly following the photolysis of OCIO in the gas phase. However, the final outcome of the near-ultraviolet photolysis of OCIO by reactions (4) and (5), namely formation of O₂ and Cl, is measurable in the laboratory and of importance in connecting OCIO to the stratospheric ozone depletion. Here we show that laboratory evidence indicates that OCIO near-ultraviolet photolysis involves just such a photoisomerization mechanism. We also discuss the spectroscopic and photochemical data available regarding the possible role of photoisomerization in the photochemistry of OCIO.

The spectroscopic information available can be used to obtain accurate excited-state structures and establish the compatibility of the proposed photochemistry with the observed changes in bond length and bond angle that occur on excitation. Specifically, activity in the asymmetric bond stretch (ν_3 in the OCIO spectrum) is needed to break a Cl-O bond. Conversely, a decreased bond angle reflected in activity in the bending mode (ν_2) in the excited state, provides the geometry change required for photoisomerization.

The near-ultraviolet spectrum of OCIO extends from 21,000 to 37,000 cm⁻¹ (270-476 nm), the observed features being attributed to a transition from the ²B₁ ground state to ²A₂ excited electronic state¹⁴⁻¹⁶. In this transition^{16,17}, the Cl-O bond length increases from 1.47 Å in the ²B₁ state to 1.619 Å in the ²A₂ excited state, and the bond angle decreases substantially from 117.6° (²B₁) to 107.0° (²A₂) as shown in Fig. 1.

The excited-state structure and vibronic lifetimes are significantly perturbed by the efficient photochemistry occurring

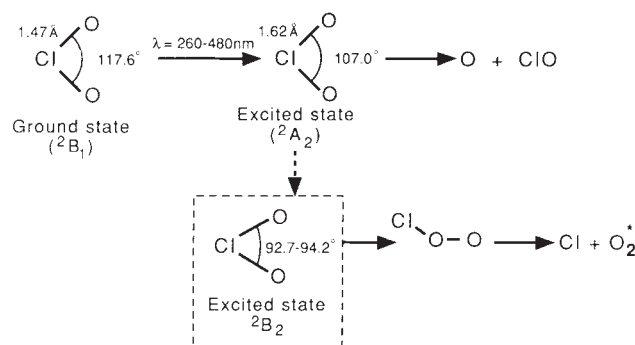


FIG. 1 Schematic diagram of the suggested photochemistry of OCIO.

on excitation of OCIO in the near-ultraviolet. The photoreactivity at this energy is presumed to be linked to coupling of the 2A_2 state with the symmetry-forbidden 2B_2 state¹⁷⁻¹⁹. The latter has eluded direct spectroscopic observation, but is calculated¹⁷ to have a minimum bond angle ranging from 94.2° to 92.7°, and consequently to favour photoisomerization¹⁷ (see Fig. 1).

In principle, examination of a structured spectrum such as that of OCIO can provide the desired information on the excited state. At room temperature, however, spectra are too congested for such an analysis because of the thermal populations of many ground-state rotational levels, all of which are involved in the electronic transition. As a result, the vibronic linewidths are dominated by thermal effects that exceed the natural linewidths and therefore preclude obtaining the excited-state lifetime.

We have re-examined the spectroscopy and photochemistry of OCIO free of inhomogeneous effects²⁰ by preparing samples in supersonic molecular jets. The jet-cooled A (2A_2) ← X (2B_1) absorption spectrum of OCIO has recently²⁰ been obtained in our laboratory with an apparatus that combines a high-resolution Fourier Transform ultraviolet spectrometer with a high-throughput molecular jet. Figure 2 shows a portion of the spectrum resulting from the transition to the 2A_2 state of OCIO. In the absence of thermal effects, the Heisenberg uncertainty principle establishes an inverse relationship between vibronic linewidths and excited-state lifetimes. Thus the cold spectrum allows dynamical information to be obtained. We can draw two conclusions from our spectra. First, all vibronic linewidths increase with energy. The corresponding lifetimes decrease from ~100 ps at 23,000 cm⁻¹ (~435 nm) to ~1 ps at 33,000 cm⁻¹ (~303 nm), reflecting the timescale of photochemical events. Second, comparing the lifetimes of the three vibronic modes of OCIO at approximately the same energy indicates that both the asymmetric stretch and the bend are involved in the photoreactivity of this molecule excited in the near-ultraviolet. As a consequence of photodissociation, the asymmetric stretch displays very short lifetimes even at low energy; the $\nu_3=4$ level, for example, becomes so broad that it disappears into the noise (having a linewidth of 100 cm⁻¹ at ~30,000 cm⁻¹). At the same time, progressions involving the bending mode (ν_2) broaden substantially at approximately the same energy, supporting the conjecture that ν_2 is involved in coupling the 2A_2 to the 2B_2 state, which is expected (in light of the small bond angle¹⁷) to favour photoisomerization.

In summary, the spectroscopic evidence is consistent with both photodissociation and photoisomerization occurring on the near-ultraviolet excited electronic surface.

The photochemistry of OCIO has been investigated extensively using flash photolysis and laser techniques²²⁻²⁵. Most authors agree that OCIO dissociates^{19,21-24} into ClO and O (reaction (2)), but some^{19,24} have suggested alternative pathways (reactions (4) and (5)), leading to Cl and O₂^{*}. Here, O₂^{*} could

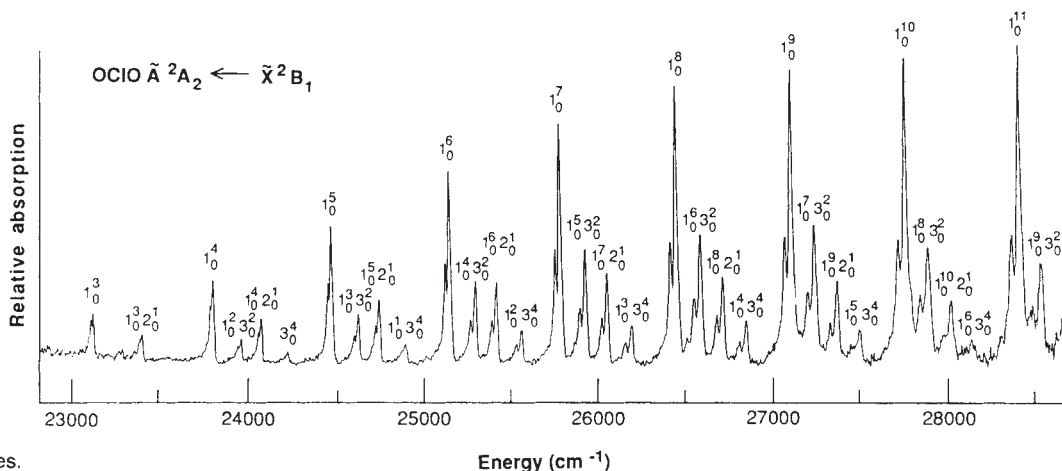
be, on the basis of energy, the $a^1\Delta_g$, the $b^1\Sigma_g$ excited states or vibrationally hot levels of the ground state ${}^3\Sigma_g$. To date, the primary photochemistry of OCIO remains controversial, as experimental studies are complicated by rapid secondary reactions that obscure the primary events.

To elucidate the primary photochemistry of OCIO in the gas phase, we are reinvestigating this problem by preparing the molecule free of collisions in supersonic expansions. The detection is based on resonance-enhanced multiphoton ionization/mass spectrometry (REMPI/MS). We measure strong, wavelength-dependent ClO⁺ and Cl⁺ mass signals following OCIO photolysis. Molecular oxygen has an extremely low multiphoton ionization (MPI) cross-section²⁶, making it difficult to detect in our system. The Cl⁺ action spectrum appears between 360 and 363 nm. It is likely that this signal reflects the coupling of OCIO and ClOO, or the REMPI structure of the isomerized parent molecule, ClOO. REMPI of other precursors such as OCIO or ClO can be excluded. Our interpretation implies photoisomerization of OCIO, followed by dissociation to produce atomic chlorine which is then probed by the MPI process. This finding is consistent with Gole's analysis that isomerization of OCIO occurs by the 2B_2 state (ref. 17), as seen in the matrix experiment of Arkell and Schwager, most efficiently at ~365 nm (ref. 27). The relative yield of the neutral photolysis products ClO, O₂ and Cl cannot be obtained from the data at this stage because of the unknown and very different MPI cross-sections which govern the signals detected. These results suggest, however, that excitation in the near-ultraviolet of isolated OCIO in the gas phase leads to both photodissociation to form ClO and to photoisomerization which ultimately leads to formation of O₂ and Cl.

The involvement of photoisomerization is compelling in light of the matrix photochemistry^{17,27,28} of OCIO. Photolysis of the symmetric chlorine dioxide in the near-ultraviolet electronic state, in argon, nitrogen or sulphuric-acid matrices, leads exclusively to formation of ClOO. Quantum chemical calculations¹⁷ explored the nature of the isomerization process and a mechanism was proposed for the concerted conversion of OCIO to ClOO. Photoisomerization is expected¹⁷ to occur by coupling of the 2A_2 state to the 2B_2 state, as already mentioned (see Fig. 1). The efficiency of this photochemical process is governed by both the position of the curve crossing between the two electronic surfaces and the magnitude of the coupling between them.

We suggest a new model for the near-ultraviolet primary photochemistry of OCIO vapour that is consistent with all the spectroscopic and photochemical data. We propose that two photochemical reactions take place, namely photodissociation and photoisomerization, as shown in Fig. 1. This model can explain the differences between the gas and condensed-phase photochemistry as it proposes that the relative quantum yields for photodissociation and photoisomerization vary with energy

FIG. 2 A portion of the jet-cooled absorption spectrum of the A ← X transition of OCIO, recorded at 5-cm⁻¹ resolution under conditions of 3% OCIO in He at a stagnation pressure of 2 atm. The splitting in each band is caused by ³⁵Cl and ³⁷Cl isotopes. All of the vibrational modes are observed in this spectrum. The most intense progression has been assigned to excitation of the symmetric stretching mode (ν_1), whereas the less intense progressions involve the bending (ν_2) and the asymmetric stretching (ν_3) modes.



and with external perturbations. This explanation is substantiated by our extensive results on other photoreactive systems²⁹⁻³⁴ which established that intermolecular interactions can shift the position of the curve crossings of the electronic surfaces responsible for chemistry and, consequently, greatly perturb the photo-product yields.

OCIO has been measured in the Antarctic stratosphere^{35,36} using several of the spectroscopic transitions depicted in Fig. 2. Both twilight and nighttime measurements have been carried out. Modelling studies based on twilight measurements³⁶ indicate that the daytime column abundance of this molecule in Antarctica is $\sim 2-6 \times 10^{12}$ molecules cm^{-2} during the fully sunlit portion of the day (for example, for solar zenith angles $\leq 85^\circ$) for much of September. The extrapolation from twilight to daytime OCIO column abundances depends on the solar-zenith-angle dependence of the OCIO photolysis rate, which is well known³⁶. The fast photolysis rate implies that the daytime OCIO distribution is in a photochemical steady state, with the production from reaction (2) being balanced by the loss that is due to photolysis. Adopting a photolysis rate³⁶ of $\sim 5 \times 10^{-2} \text{ s}^{-1}$ along with the daytime column abundances indicated above leads to an estimated column integrated rate of OCIO photolysis of $\sim 1-3.0 \times 10^{11}$ molecules $\text{cm}^{-2} \text{ s}^{-1}$. If photolysis of OCIO proceeds entirely through photoisomerization (leading to ozone loss), then the corresponding contribution to the column-integrated loss of Antarctic ozone would be $\sim 0.1-0.3\%$ per day. The observed rate of Antarctic total ozone decline is $\sim 1-2\%$ per day, so that this mechanism could contribute an upper limit of $\sim 5-30\%$ of the Antarctic total ozone loss. As noted earlier, the photoisomerization of OCIO is quantitative in the condensed phase²⁷, whereas the quantum yields for the photochemical reactions of this molecule in the gas phase have yet to be quantitatively established. The fractional contribution of this mechanism to Arctic ozone destruction is likely to be far greater than in Antarctica, and hence could be quite significant even if the branching ratio for photoisomerization is less favourable in the gas phase than in matrices. $\square$

Natural oscillations of small raindrops

Kenneth V. Beard*[†], Harry T. Ochs III*
& Rodney J. Kubesh*[†]

* Climate and Meteorology Section, Illinois State Water Survey and

[†] Department of Atmospheric Sciences, University of Illinois, 105 South Gregory Avenue, Urbana, Illinois 61801, USA

REMOTE sensing of the rate of rainfall and the characteristics of hydrometeors with polarimetric radars is based on the assumption that the eccentricity of raindrops increases predictably with drop size¹⁻³. Field observations, however, have shown that a particular raindrop size can have a wide range of shapes^{4,5} because of oscillations thought to be caused by drop collisions. Here we report observations of the shapes of small raindrops in circumstances where oscillations might occur sympathetically with eddy shedding in the drop's wake. The analysis of oscillations was aided by measurements of oscillation frequency, based on variations in rainbow angle and comparisons of mode energies. Our observations agree with quiescent shapes only for drops < 1 mm in diameter. For 1–1.5 mm we discovered significant shape variability of distinct types corresponding to drop oscillations in axisymmetric and transverse modes. Our finding of reduced average distortion explains the mysterious shift away from quiescent distortion for small raindrops, postulated from radar observations⁶, and indicates the existence of an important class of raindrop oscillations—promoted not by collisions but by intrinsic aerodynamic forces.

The shapes of raindrops are calculated using a force balance between surface tension and the aerodynamic pressure distribution for a uniform fall speed⁷⁻⁹. The quiescent shape of a small raindrop is essentially an oblate spheroid, but a larger raindrop tends toward the shape of a sessile drop with a pronounced flattening of the base. The calculated shapes of raindrops show good agreement with wind-tunnel measurements of quiescent drop shape for sizes > 2 mm diameter⁷⁻¹⁰. Field observations of moderate to large raindrops, however, reveal a wide variation about the quiescent shape^{4,5}. Thus standard assumptions of raindrop shape are being re-examined for radar measurements of rain^{5,6}.

Nonquiescent shapes are thought to be simply the result of oscillations. For large raindrops, collisions with the more numerous drizzle drops are frequent enough to maintain oscillations, but for small raindrops another mechanism is required because collisions are too infrequent¹¹. One possibility is that drop oscillations are caused by the feedback of intrinsic aerodynamic forces, for example by eddies detaching from drops of ~ 1 mm diameter where the shedding frequency matches the primary oscillation frequency¹². This resonance phenomenon has never been observed because 1-mm drops do not fall straight down and therefore cannot be contained within a wind tunnel or an acceleration column. The sideways drift occurs when the flow around a falling spheroid becomes asymmetric.

For this study we made the first observations of nonquiescent shapes for small raindrop sizes. Our approach minimized the sideways drift by generating drops at terminal velocity in still air by using an orifice-jet drop generator¹³; we were then able to study their natural behaviour in an experiment 4 m in height, a distance sufficient for initial oscillations produced during drop generation to decay by viscous dissipation to an immeasurable amplitude. The drop shape was obtained from stroboscopic photos of drop silhouettes, whereas the oscillation frequency was determined from separate streak photos taken with a xenon arc lamp. Measurements of shape were made from photographic negatives by using a low-power microscope with an eyepiece micrometer to obtain the maximum vertical extent (v), maximum horizontal extent (h) and axis ratio ($\alpha = v/h$).

Received 9 June; accepted 6 October 1989.

- Farman, J. C., Gardiner, B. G. & Shanklin, J. D. *Nature* **315**, 207–210 (1985).
- Stolarski, R. S. *et al.* *Nature* **322**, 808–811 (1986).
- Hofmann, D. J., Harder, J. W., Rolf, S. R. & Rosen, J. M. *Nature* **326**, 59–62 (1987).
- Gardiner, B. G. & Shanklin, J. D. *Geophys. Res. Lett.* **13**, 1199–1201 (1986).
- Solomon, S., Garcia, R. R., Rowland, F. S. & Wuebbles, D. J. *Nature* **321**, 755–758 (1986).
- McElroy, M. B., Salawitch, R. J., Wofsy, S. C. & Logan, J. A. *Nature* **321**, 759–762 (1986).
- Crutzen, P. J. & Arnold, F. *Nature* **324**, 651–655 (1986).
- Barrett, J. W. *et al.* *Nature* **336**, 455–458 (1988).
- Anderson, J. G., Brune, W. R. & Chan, R. J. *Geophys. Res.* **94**, 11480–11520 (1989).
- Jones, R. L. *et al.* *J. geophys. Res.* **94**, 11529–11558 (1989).
- Sander, S., Friedl, R. & Yung, Y. *Science* **245**, 1095–1097 (1989).
- Pyle, J. A. *Geophys. Res. Lett.* **13**, 1320–1322 (1986).
- Tuck, A. F. *J. geophys. Res.* **94**, 11687–11737 (1989).
- Coon, J. B. & Ortiz, E. *J. molec. Spectrosc.* **1**, 81–94 (1957).
- Wahner, A., Tyndall, G. S. & Ravishankara, A. R. *J. phys. Chem.* **91**, 2734–2738 (1987).
- Brand, J. C. D., Redding, R. W. & Richardson, A. W. *J. molec. Spectrosc.* **34**, 399–414 (1970).
- Gole, J. L. *J. Phys. Chem.* **84**, 1333–1340 (1980).
- Michielsen, S. *et al.* *J. chem. Phys.* **74**, 3089–3101 (1981).
- McDonald, P. A. & Innes, K. K. *Chem. Phys. Lett.* **59**, 562–566 (1978).
- Richard, E. C., Wickham-Jones, C. T. & Vaida, V. *J. phys. Chem.* **93**, 6346–6350 (1989).
- Lipscomb, F. J., Norrish, R. G. W. & Thrush, B. A. *Proc. R. Soc. A* **233**, 455–464 (1956).
- Bethune, D. S., Schell-Sorokin, A. J., Lankard, J. R., Loy, M. M. T. & Sorokin, P. P. in *Advances in Laser Spectroscopy*, Vol. 2 (eds Garetz, B. A. & Lombardi, J. R.) (Wiley Chichester, 1983).
- Glowina, J. H., Misewich, J. & Sorokin, P. P. in *Supercontinuum Laser Sources* (ed. Alfano, R. R.) (Springer, Berlin, in the press).
- Hayman, G. D. & Cox, R. A. *Chem. Phys. Lett.* **155**, 1–7 (1989).
- Colussi, A. J., Redmond, R. W. & Scialano, J. C. *J. phys. Chem.* **93**, 4783–4785 (1989).
- Hodges, R. V., Lee, L. C. & Moseley, J. T. *Int. J. Mass Spectrom. Ion Phys.* **39**, 133–143 (1981).
- Arkell, A. & Schwager, I. *J. Am. chem. Soc.* **89**, 5999–6006 (1967).
- Adrian, F. J., Bohandy, J. & Kim, B. F. *J. chem. Phys.* **85**, 2692–2698 (1986).
- Donaldson, D. J., Gaines, G. A. & Vaida, V. *J. phys. Chem.* **92**, 2766–2769 (1988).
- Sapers, S. P., Vaida, V. & Naaman, R. *J. chem. Phys.* **88**, 3638–3645 (1988).
- Donaldson, D. J., Richard, E. C., Strickler, S. J. & Vaida, V. *J. phys. Chem.* **92**, 5514–5517 (1988).
- Donaldson, D. J., Child, M. S. & Vaida, V. *J. chem. Phys.* **88**, 7410–7417 (1988).
- Vaida, V., Donaldson, D. J., Sapers, S. P., Naaman, R. & Child, M. S. *J. phys. Chem.* **93**, 513–520 (1989).
- Prinslow, D. A. & Vaida, V. *J. phys. Chem.* **93**, 1836–1840 (1989).
- Solomon, S., Mount, G. H., Sanders, R. W. & Schmeltkeop, A. L. *J. geophys. Res.* **92**, 8329–8338 (1987).
- Solomon, S., Sanders, R. W., Carrol, M. A. & Schmeltkeop, A. L. *J. geophys. Res.* **94**, 11393–11404 (1989).

ACKNOWLEDGEMENTS. This work was supported by the Petroleum Research Fund and the National Science Foundation. E.R. acknowledges financial support by Deutsche Forschungsgemeinschaft.

Figure 1 shows the results from our measurements of axis ratio as a function of drop diameter d . For the smallest drops ($d \leq 1$ mm), our observations of axis ratios agree with the theoretical quiescent shape^{7,9} within experimental uncertainty. The previous experimental data at small sizes¹⁰ shows approximately the same scatter but seems to have a systematic error. In the present experiment, such errors were carefully evaluated and minimized using glass spheres to compare axis ratios obtained using the photographic method with those from direct measurements.

The measured axis ratios show a large two-sided scatter for $d \geq 1.4$ mm and a moderate one-sided scatter, above the theoretical curve, for $d \approx 1.0$ –1.3 mm. The large two-sided scatter can be explained by axisymmetric mode oscillations of oblate- and prolate-like shapes having a two-sided axis-ratio variation, as Fig. 2 shows. Also, our measurements of the axes for the largest sizes revealed the characteristic opposing variation in v and h . The observed one-sided scatter for $d \approx 1.0$ –1.3 mm is consistent

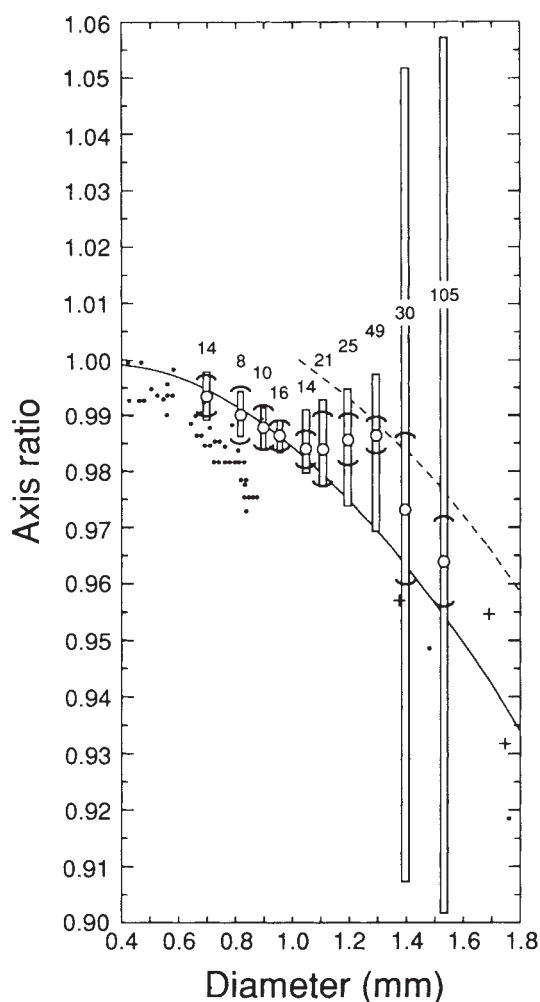


FIG. 1 Measured axis ratios as a function of drop diameter. The vertical bars show the range of measured values, including scatter and estimated (r.m.s.) experimental uncertainty. The uncertainty decreases with size at a fixed magnification because a measurement of axis ratio becomes more precise. Also displayed are mean axis ratios (circles), 95% confidence intervals (arcs) and number of measurements. The width of the bars indicates uncertainty from minor changes in drop size during (or between) experiments. About 4,000 photographs were needed to obtain 292 axis ratios, because of a preponderance of out-of-focus drops (horizontal scatter in fall trajectories was much larger than depth of field). The theoretical quiescent axis ratio is the solid curve⁷⁻⁹ and previous wind-tunnel measurements are dots¹⁰ and crosses⁷. The dashed curve shows the postulated shift in axis ratio based on polarization radar measurements in rain⁶.

with transverse-mode oscillations of the lowest harmonic. As Fig. 2 shows, the maximum vertical extent increases more than the maximum horizontal extent for various side views so that observations of v/h should yield a one-sided variation above the quiescent axis ratio.

Further support for these postulated modes is the consistency of oscillation energies despite the different axis-ratio amplitudes in Fig. 1. Energies are comparable if the transverse mode is assumed for $d \approx 1.0$ –1.3 mm and the axisymmetric mode is assumed for $d \geq 1.4$ mm (~ 0.2 – 0.4% of the surface energy). By contrast, energies differ by a factor of 20 if the reverse is assumed (axisymmetric mode at the smaller sizes and transverse mode at the larger sizes).

The onset of raindrop oscillations at $d \approx 1.0$ mm is close to the onset of eddy shedding for falling drops and spheres (Fig. 3). In addition, the observed onset of sideways drift also occurs just above $d = 1.0$ mm. For drops of $d = 1.0$ –1.3 mm the match between oscillation and eddy frequencies is best at the lowest harmonic ($n = 2$). The two largest sizes have a better match with the two higher harmonics ($n = 3$ and 4). Although different harmonics have distinct shapes, it was difficult to distinguish among them from the drop photos because the single view of the profile could correspond to more than one mode, and the distortions were hard to identify at these small amplitudes.

Another distinguishing characteristic of each harmonic is its oscillation frequency, for example $n = 3$ has a frequency about twice that frequency of $n = 2$ (Fig. 3). Oscillation frequencies were measured by recording drop-fall streaks from a xenon light by using the modulation of the rainbow (Fig. 4). Large-amplitude oscillations produce dashed streaks in back-scattered light¹⁴, as can be seen by shining a floodlight at rain or water dripping from a tree branch. Such streaks occur when the angle Ψ changes around the primary rainbow angle as light is deviated from the bow into the dark band. Our streak photos revealed dashes of spacing λ corresponding to the frequency of the fundamental harmonic f_2 ($f \approx f_2$), using the known terminal fall

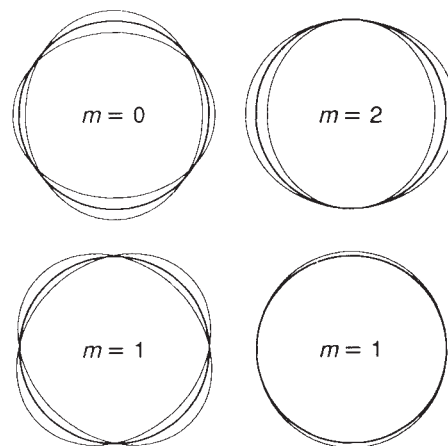


FIG. 2 Horizontal views (thin lines) for fundamental spherical harmonic ($n = 2$) and unperturbed spheres (thick lines). The three modes ($m = 0, 1, 2$) of equal oscillation energy (1% of surface energy) correspond to associated Legendre functions. The axisymmetric mode ($m = 0$) has oscillations of oblate- and prolate-like shapes with a two-sided axis-ratio variation about unity. The transverse mode ($m = 1$) is shown for two orientations (side views), 90° apart in the horizontal plane. The maximum vertical extent (v) increases more than the maximum horizontal extent (h) for various azimuthal views so that observations of v/h yield a one-sided variation above unity. The horizontal mode ($m = 2$) oscillates in the horizontal plane with major and minor axes 90° apart. The vertical axis remains fixed whereas the horizontal axis is larger or smaller than the vertical axis, producing axis ratios with a two-sided variation about unity. All three modes have been observed for water drops falling in air with oscillations about a quiescent raindrop shape rather than a sphere²⁰.

speed U and $f = U/\lambda$, with an average experimental uncertainty of only 3% in the measurement of f .

The streak observations substantiate our conclusions based on the behaviour of the scatter in axis ratio: small raindrops oscillate at the fundamental harmonic in the transverse and axisymmetric modes. At the onset of eddy shedding the fundamental harmonic is excited, as might be expected from the match in frequencies (Fig. 3). Oscillations occur in the transverse mode perhaps as a response to pressure changes caused by eddies detaching from alternate sides of the upper pole. At larger sizes the oscillation response to eddy shedding is the axisymmetric mode of the fundamental harmonic, despite the poor match of $n = 2$ with the shedding frequencies¹⁵⁻¹⁹.

The cause of the broad coupling between the fundamental mode and eddy shedding is unknown. One possibility is that, once oscillations become large enough, shedding is triggered by shape changes, thereby shifting the shedding frequency to the oscillation frequency. Another possibility is a subharmonic coupling of eddies to the oscillations. If pressure patterns were alternatively in and out of phase with the distortion, this could explain why the transverse mode is not excited. It is unclear how the forcing pattern would best match the axisymmetric mode or why the fundamental is preferred to the next higher harmonic.

An indication that the observed strong-coupling phenomenon may extend to much larger raindrop sizes is apparent from the consistency of our axis ratios with those from the differential reflectivity, Z_{DR} , measurements in rain⁶ (Fig. 1). As Z_{DR} is proportional to the logarithm of the ratio of backscattered powers for radar pulses having vertical and horizontal polarizations, it is directly related to the axis ratio of raindrops. The variation in axis ratio given by the dashed curve in Fig. 1 was devised to rectify values of Z_{DR} measured with a high-resolution radar (Chilbolton, UK) with values calculated from observed raindrop size distributions. This postulated shift in axis ratio, which was applied as an average effect from 1 to 3 mm diameter, is a likely consequence of raindrop oscillations promoted by eddy shedding, at least for small raindrops. The occurrence of raindrop oscillations for $d = 2-3$ mm is also supported by recent

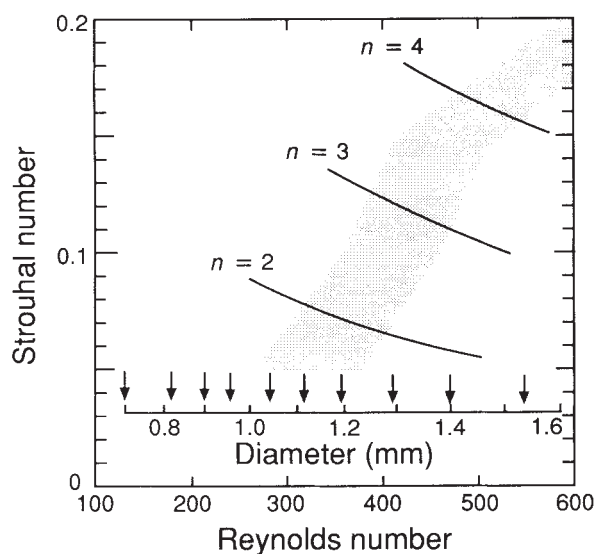


FIG. 3 Dimensionless frequency (Strouhal number = fd/U) as a function of water-drop diameter and dimensionless flow parameter (Reynolds number = dU/ν). The oscillation frequencies for water drops in air are given by curves for harmonics $n=2, 3$ and 4 . Previous measurements of eddy shedding frequencies are shown by the stippled area for drops^{15,16} and rigid spheres¹⁷ falling in liquids in a region with Strouhal number <0.15 and for supported spheres^{18,19} in a region with Strouhal number >0.15 . These data are for nonoscillating drops and rigid spheres. The arrows show sizes for drops falling in air in the present experiment.

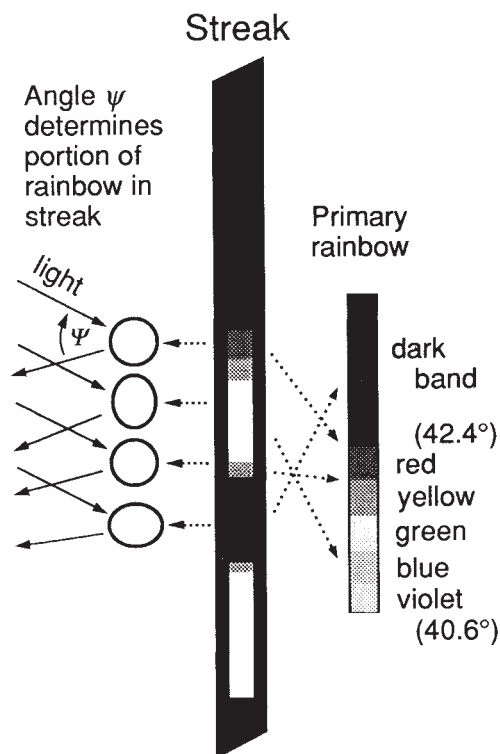


FIG. 4 Schematic of modulated fall streak showing the associated drop oscillation and rainbow. The oscillation frequency was measured from the spacing of dashes in photos of the fall streaks. The dotted arrows show how portions of the fall streak correspond to the origin of colour in the rainbow, and how drop distortions yield a change in angle Ψ . The angles are shown for red and violet. Colours are weak, except near red, because the width of the light allows incident rays over a range of angles, thereby mixing the adjacent colours.

airborne observations of raindrop shape⁵. Such oscillations of moderate-size raindrops probably originate from collisions¹¹, but may persist because of intrinsic aerodynamic forces, for example from feedback of the varying aerodynamic pressure or drag, opposing the viscous damping of oscillations.

To obtain further data on larger drops, we plan to extend our measurements of natural shapes up to 4 mm in diameter by using our experimental methods in a seven-story experiment. A careful analysis of shape variations, oscillation amplitudes and oscillation frequencies should uncover more of the mysterious physics that underlies raindrop oscillations, and provide valuable information for understanding the scattering of microwave radiation in rain. □

Received 17 July; accepted 10 October 1989.

1. Seliga, T. A. & Bringi, V. N. *J. appl. Met.* **15**, 69-76 (1976).
2. Hall, M. P. M., Cherry, S. M., Goddard, J. W. F. & Kennedy, G. R. *Nature* **285**, 195-198 (1980).
3. Illingworth, A. *Nature* **336**, 754-756 (1988).
4. Jones, D. M. A. *J. Met.* **16**, 504-510 (1959).
5. Chandrasekar, V., Cooper, W. A. & Bringi, V. N. *J. Atmos. Sci.* **45**, 1323-1333 (1988).
6. Goddard, J. W. F. & Cherry, S. M. *Radio Sci.* **19**, 201-208 (1984).
7. Pruppacher, H. R. & Pitter, R. L. *J. Atmos. Sci.* **28**, 86-94 (1971).
8. Beard, K. V. & Chuang, C. *J. Atmos. Sci.* **44**, 1509-1524 (1987).
9. Beard, K. V., Feng, J. Q. & Chuang, C. *J. Atmos. Sci.* **46**, 2404-2418 (1989).
10. Pruppacher, H. R. & Beard, K. V. *J. R. Met. Soc.* **96**, 247-256 (1970).
11. Beard, K. V., Johnson, D. B. & Jameson, A. R. *J. Atmos. Sci.* **40**, 455-462 (1983).
12. Gunn, R. *J. geophys. Res.* **54**, 383-385 (1949).
13. Czys, R. R. & Ochs, H. T. *J. Atmos. Sci.* **45**, 3161-3168 (1988).
14. Blanchard, D. C. *J. Atmos. Sci.* **19**, 119-120 (1962).
15. Magarvey, R. H. & Bishop, R. L. *Can. J. Phys.* **39**, 1418-1423 (1961).
16. Magarvey, R. H. & Bishop, R. L. *Phys. Fluids* **4**, 800-805 (1961).
17. Goldburg, A. & Florsheim, B. H. *Phys. Fluids* **9**, 45-50 (1966).
18. Achenbach, E. *J. Fluid Mech.* **62**, 209-221 (1974).
19. Möller, W. *Phys. Z.* **39**, 57-80 (1938).
20. Beard, K. V. *J. Atmos. Sci.* **41**, 1765-1774 (1984).

ACKNOWLEDGEMENTS. This research was supported by the Physical Meteorology Program, Division of Atmospheric Sciences, U.S. National Science Foundation.

Simulation of the regional climatic impact of Amazon deforestation

J. Lean & D. A. Warrilow*

Dynamical Climatology Branch, Meteorological Office, Bracknell, Berkshire RG12 2SZ, UK

* Present Address: Global Atmosphere Division, Department of Environment, London SW1P 3PY, UK

THE Amazon basin contains about half of Earth's Tropical forest¹. Population pressure and subsequent demands for crop production, timber and firewood have led to rapid deforestation. Quantitative estimates of the rate of deforestation from analysis of Landsat observations indicate that rates are increasing exponentially in many regions^{2,3}, but the precise figures are not known⁴. Removal of the protection provided by natural cover can lead to soil erosion, disturbance of the ecosystem and reduction in species diversity⁵. Here we report results from a three-year simulation, using a general circulation model, in which we replace Amazon tropical forest and savannah with pasture. The simulated local climate response was dominated by a weakened hydrological cycle, with less precipitation and evaporation and an increase in surface temperature. The reductions in precipitation and evaporation were mostly caused by changes in surface roughness and albedo: decreased roughness dominated the reduction in evaporation (and the increase in temperature), whereas the increased albedo was the main cause of a decrease in the moisture flux convergence (measured as the difference between precipitation and evaporation) contributing to the decrease in precipitation.

Within the Amazon basin, it is estimated that 50% of the rainfall is derived from local evaporation⁶, and it has been suggested that a significant reduction in tropical forest might reduce evaporation and rainfall⁷. Recently, various studies concerning the impact of surface changes on climate have been carried out using three-dimensional atmospheric global climate models (AGCMs). A review of tropical deforestation experiments⁸ shows that substantial differences exist between the schemes for representing physical processes (parameterizations) and between methods used to simulate land transformations in AGCMs. As a result AGCMs have not presented a uniform picture of the potential effects of deforestation on climate. The use of rather simple land surface schemes and the lack of grid-scale data with which to calibrate the models make the applicability of the results uncertain. Some common points, however, have emerged from these simulations. First, albedo, roughness and surface hydrology are the most important variables for the representation of land transformation. Second, results are sensitive to the parameterization processes in the model, to the choice of vegetation and soil parameters and the soil-moisture field as input data, and the climatology of the AGCM. Over the past few years more realistic land surface parameterizations have been included in various AGCMs including that of the UK Meteorological Office.

Here we describe experiments that were carried out using the $2\frac{1}{2}^\circ \times 3\frac{3}{4}^\circ$ resolution AGCM described in ref. 9, but including a more complex representation of interactive cloud processes¹⁰. The land surface parameterization (D.A.W., A. B. Sangster and A. Slingo, unpublished results) includes a four-layer soil temperature scheme. Interception of precipitation by a vegetation canopy is represented, allowing zero resistance to re-evaporation and potentially a more realistic partitioning of energy between latent- and sensible-heat fluxes. The hydrological scheme diagnoses surface runoff from the excess of rainfall over infiltration capacity (accounting for the spatial variability of rainfall statistically), and subsurface runoff from the gravitational drainage of soil moisture¹¹. Geographical variation of land surface and soil

types is based on the work of Wilson and Henderson-Sellers¹². Their original $1^\circ \times 1^\circ$ -resolution primary- and secondary-vegetation datasets were used to create spatially varying datasets of 17 soil and vegetation parameters which were interpolated onto our $2\frac{1}{2}^\circ \times 3\frac{3}{4}^\circ$ model grid¹³.

We ran three-year control and deforestation experiments, starting at the end of June. All grid boxes in South America north of 30° S defined as either tropical forest or savannah on the $1^\circ \times 1^\circ$ grid were replaced by tropical pasture. Soil characteristics except surface infiltration capacity, were left unaltered. We give the results for an area in the southern part of Amazonia where the largest change in land surface characteristics occurred (see Fig. 1).

We compare the control simulation with observation in Fig. 1 and Table 1. The model simulates accurately the spatial distribution of rainfall for both the Southern Hemisphere winter and summer, although the dry areas are more extensive in winter over the southern Amazon and rainfall is slightly excessive in summer over the basin. Looking particularly at the region concerned, Fig. 2a shows this slight amplification of the seasonal cycle in the control rainfall simulation. The annual mean values of other variables, compared in Table 1, show that the control simulation is quite realistic. We note that it is extremely difficult to obtain accurate estimates of such quantities because observations are too sparse to provide grid-square means. Runoff is

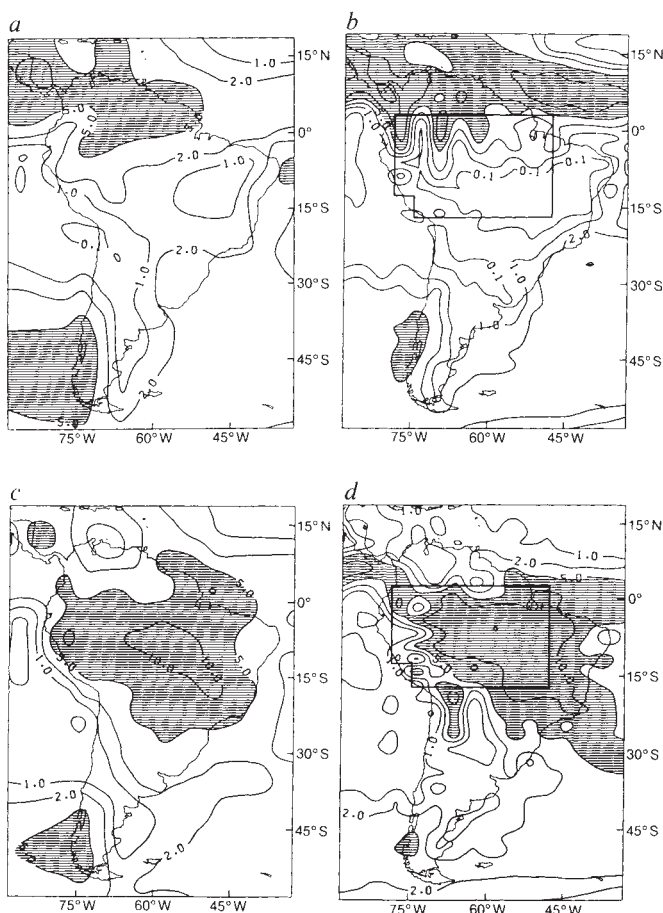


FIG. 1 The section within the box shown in *b* and *d* represents the area selected for averaging purposes for all results quoted and represents the area having the largest change in land surface characteristics in the southern region north of 30° S. *a*, Observed²¹ precipitation pattern over South America for the Southern Hemisphere winter (June, July and August). Contours are at 0.1, 1.0, 2.0, 5.0, 10.0 and 20.0 mm d⁻¹ and shading is above 5 mm d⁻¹. *b*, Simulated precipitation pattern of the control experiment over South America for the Southern Hemisphere winter averaged over three years. *c*, As for *a* but for the Southern Hemisphere summer (December, January and February). *d*, As for *b* but for the Southern Hemisphere summer.

TABLE 1 Summary of surface variables

Surface Variable	C*	D†	O‡
Evaporation (mm d^{-1})	3.12	2.27 (−27.2%)	3.34 (ref. 21)
Precipitation (mm d^{-1})	6.60	5.26 (−20.3%)	5.26 (ref. 22)
Soil moisture (cm)	16.13	6.66 (−58.7%)	
Runoff (mm d^{-1})	3.40	3.0 (−11.9%)	2.76 (ref. 23)
Net radiation (W m^{-2})	147.3	126.0 (−14.5%)	
Temperature ($^{\circ}\text{C}$)	23.6	26.0 (+2.4 $^{\circ}\text{C}$)	24.0 (ref. 24)
Sensible heat (W m^{-2})	57.2	60.2 (+5.2%)	
Bowen Ratio	0.85	1.5 (+76.5%)	

* Control simulation, averaged over three years.

† Deforested simulation, averaged over three years.

‡ Observations.

only quoted for a particular basin and is not strictly comparable to that from the region shown. Evaporation is based on only a few direct observations and empirical calculations.

Comparison of surface quantities for a single model grid square with data¹⁴ for a site in central Amazonia indicates that modelled total evaporation differs by 15% from the observed, and that the modelled canopy evaporation is 71% compared with the observed 25% of total evaporation. Overestimation of canopy evaporation is probably present in other land surface schemes and this may be due¹⁵ to the extension of a single-point description of the rainfall interception process to the grid-scale area in a region where convective rainfall events dominate.

Table 1 also summarizes the effect of deforestation on local climate for a variety of surface variables from each experiment. The hydrological cycle is significantly weakened by deforestation, with a reduction in evaporation, rainfall and runoff. Surface runoff is reduced as the decrease in intensity and frequency of rainfall events predominates over an opposing tendency for increased runoff that is the result of a reduction in infiltration capacity. The decrease in available soil moisture occurs as a result of the marked reduction in root depth (from 127.4 cm to 61.9 cm). Pasture reflects more solar radiation than does forest, which is a very efficient absorber and scatterer of short-wavelength radiation. Consequently, the net energy absorbed by the surface was diminished, with the mean albedo increased from 0.136 to 0.188. By itself, a reduction in net

absorption of radiation at the surface would result in a temperature decrease but this is offset by a reduction in evaporative cooling and so there is an overall rise in temperature. The decrease in the latent-heat flux was due to the reduction in roughness (from 0.79 m to 0.04 m) and, at some times of the year, the water availability; this decrease is primarily responsible for the marked increase in the Bowen ratio (defined as the ratio of sensible to latent heat).

Figure 2 depicts the seasonal variation of precipitation and evaporation for the control simulation, the 'deforested' simulation and the difference between the two. There is a consistent decrease in precipitation and evaporation throughout the year although rainfall changes exhibit greater variability. There is also a seasonal variation in the change in evaporation with the maximum differences occurring in the winter and spring when evaporation is controlled by the limited availability of soil moisture.

To gain a better understanding of the physical mechanisms that are operating, two eight-month simulations, using the same initial data as the control and deforested experiments, were carried out to investigate the relative impact that is due to changes to albedo and roughness alone. Table 2 summarizes the differences between some of the surface variables of the control simulation and those of the experiments with increased albedo (from 0.136 to 0.188), decreased roughness (from 0.79 m to 0.04 m), and total deforestation. Precipitation within the Amazon basin originates from two sources—from the recycling of water vapour released during evapotranspiration, and from the moisture flux convergence into this area. The decrease in moisture flux convergence (measured as the difference between precipitation and evaporation) is due mainly to increased albedo although there is some contribution from the change in roughness. Increased albedo reduces the heat source, weakening ascent and boundary-layer convergence¹⁶. Reduced roughness can also decrease the moisture flux convergence as it weakens the boundary-layer convergence (associated with frictional drag near the surface) into the continental surface-pressure low¹⁷. Decreased roughness dominates the reduction in evaporation and the associated temperature increase. The forest canopy is wet for most of the year (implying low surface resistance) and so a decrease in evaporation is to be expected as roughness decreases¹⁸. Increased albedo reduces the available energy for upward turbulent transfer of latent energy. Hence, changes in both albedo and roughness make a major contribution to the overall decreases in precipitation and evaporation in the deforestation experiment.

Our results can be compared with a 13-month simulation carried out by Dickinson and Henderson-Sellers¹⁹. They incorporated a land surface scheme with a canopy in the National Center for Atmospheric Research Community Climate Model which has a grid resolution of $4.5^{\circ} \times 7.5^{\circ}$. They assumed that all of the Amazon tropical forest in South America was replaced by impoverished grassland. Both experiments show a decrease in evaporation ($\sim 16 \text{ mm month}^{-1}$ in ref. 19 (from their Fig. 11) compared with 26 mm month^{-1} here) and an increase in surface temperature (with a range of $2.0\text{--}5.0^{\circ}\text{C}$ over the 13 months in ref. 19 compared with the average of 2.4°C here). Dickinson

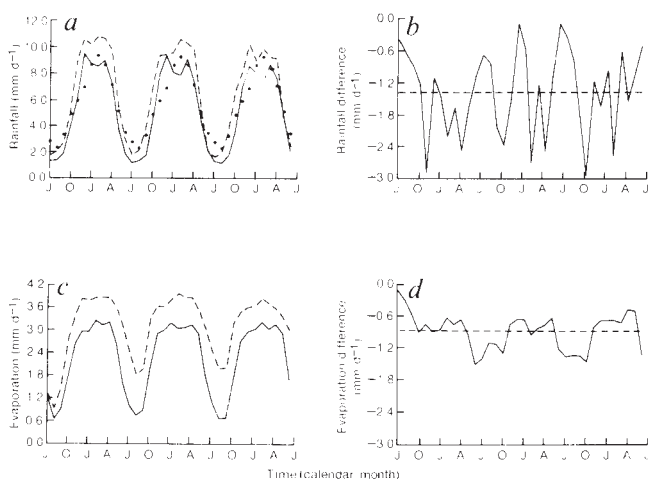


FIG. 2 *a*, Monthly mean precipitation (mm d^{-1}) averaged over the region under consideration for three years starting in July. The broken line represents the control simulation (C), the solid line the deforested simulation (D) and the dots the observed rainfall (O) (ref. 21). *b*, Simulated monthly mean difference (D-C) of precipitation (mm d^{-1}) averaged over the region under consideration for three years starting in July. The broken line represents the mean difference precipitation. *c*, As for *a* but for evaporation (mm d^{-1}). *d*, as for *b* but for evaporation (mm d^{-1}).

TABLE 2 Summary of the differences in surface variables

Surface Quantity	A-C*	R-C†	D-C‡
Evaporation (E) (mm d^{-1})	−0.20	−0.43	−0.61
Precipitation (P) (mm d^{-1})	−0.75	−0.69	−1.34
P-E (mm d^{-1})	−0.55	−0.26	−0.73
Temperature ($^{\circ}\text{C}$)	−0.10	2.24	1.98

These results are averaged over eight months.

* A, increased albedo simulation; C, control simulation.

† R, decreased roughness simulation.

‡ D, 'deforested' simulation.

and Henderson-Sellers, however, found no detectable change in rainfall, in contrast to the average decrease of 1.3 mm d^{-1} (20%) shown here.

Our results indicate a stronger impact of deforestation on the local climate than do previous GCM experiments. We note, however, that these results are highly sensitive to the formulations of the model and our ability to make firm predictions based on these results is inhibited by the lack of observations in forested and deforested areas with which to verify these schemes. More ground-based measurements of atmospheric and hydrological processes, together with global data sets of surface vegetation types derived from satellite data, could aid in this development, both as input to the model and to validate and adjust the land surface parameterizations²⁰. Nevertheless, our results indicate that deforestation can cause significant local climatic perturbations and further efforts should be made to improve our understanding of the processes involved. □

Received 22 August; accepted 13 October 1989.

1. Dickinson, R. E. (ed.) *The Geophysics of Amazonia* 3–10 (Wiley, New York, 1987).
2. Fearnside, P. M. in *The Geophysics of Amazonia* (ed. Dickinson, R. E.) 37–61 (Wiley, New York, 1987).
3. Malingreau, J. P. & Tucker, C. J. *Ambio* **17**, 49–55 (1988).
4. Neto, R. B. *Nature* **339**, 86 (1989).
5. Lavelle, P. in *The Geophysics of Amazonia* (ed. Dickinson, R. E.) 175–223 (Wiley, New York, 1987).
6. Salati, E., Marques, J. & Molion, L. C. B. *Interciencia* **3**, 200–205 (1978).
7. Salati, E. in *The Geophysics of Amazonia* (ed. Dickinson, R. E.) 273–296 (Wiley, New York, 1987).
8. Henderson-Sellers, A. in *The Geophysics of Amazonia* (ed. Dickinson, R. E.) 463–496 (Wiley, New York, 1987).
9. Slingo, A., Wilderspin, R. C. & Smith, R. N. B. *J. geophys. Res.* **94**, 2281–2301 (1989).
10. Smith, R. N. B. *Q. J. R. met. Soc.* (in the press).
11. Warrilow, D. A. in *Proc. Int. Satellite Land-Surface Climatology Proj.* 159–166 (European Space Agency, 1986).
12. Wilson, M. F. & Henderson-Sellers, A. *J. Clim.* **5**, 119–143 (1985).
13. Warrilow, D. A. & Buckley, E. *Annls Geophys.* **7**, 439–450 (1989).
14. Shuttleworth, W. J. *Proc. R. Soc.* **B233**, 321–346 (1988).
15. Shuttleworth, W. J. & Dickinson, R. E. *Q. J. R. met. Soc.* **115**, 1177–1179 (1989).
16. Charney, J. G. *Q. J. R. met. Soc.* **101**, 193–202 (1975).
17. Sud, Y. C., Shukla, J. & Mintz, Y. *J. appl. Met.* **27**, 1036–1054 (1988).
18. Rowntree, P. R. *UN. Univ. Workshop, Forests, Climate and Hydrology—Regional Impacts* (United Nations University, Tokyo, 1984).
19. Dickinson, R. E. & Henderson-Sellers, A. *Q. J. R. met. Soc.* **114**, 439–462 (1988).
20. *Int. Satellite Land-Surface Climatology Proj. Report No. 10* (eds Becker, F., Bolle, H. J. & Rowntree, P. R.) (Free University of Berlin, 1987).
21. Willmott, C. J., Rowe, C. M. & Mintz, Y. *J. Clim.* **5**, 589–606 (1985).
22. Jaeger, L. *Monatskarten Des Niederschlags Fur Die Ganze Erde. Bericht Deutsche Wetterdienst* **18**, no. 139 (1976).
23. *Discharge of Selected Rivers of the World*, Vol. 2 (UNESCO, Paris, 1971).
24. Schutz, C. & Gates, W. L. *Report for Advanced Research Projects Agency; Global Climatic Data for Surface, 800 mb, 400 mb, January (1971); July (1972); April (1973); October (1974)*; (Rand, Santa Monica).

ACKNOWLEDGEMENTS. We thank P. R. Rowntree for his guidance in the analysis.

Climate patterns revealed by pollen and oxygen isotope records of a Tyrrhenian sea core

Martine Rossignol-Strick & Nadine Planchais

Laboratoire de Palynologie, URA 327 du CNRS, Université de Montpellier II, Sciences et Techniques du Languedoc, 34095 Montpellier Cedex 2, France

THE analysis of pollen from marine cores has produced continental palaeoclimate records which have been directly correlated with the oxygen isotope record of global ice volume and regional climate^{1–3}. Here we tie the palaeoclimate of southern Europe to this global climate signal by reporting continuous pollen and $\delta^{18}\text{O}$ records of a well-dated core in the Tyrrhenian Sea for the time span of 55–9 kyr BP. These records show a strong covariation between the marine $\delta^{18}\text{O}$ reversed curve and a good terrestrial climate indicator, namely the pollen abundance of deciduous oak from the upland sub-humid Mediterranean forest of southern Italy and Sicily, which seems mainly constrained by moisture variation. The *Artemisia* and grass pollen abundances document the shift between

continental semi-desert and oceanic steppe climates. Repetitive successions between 55 and 33 kyr BP of grass, oak, *Artemisia* and *Abies* pollen abundance peaks occur in locked phase with $\delta^{18}\text{O}$ depletion events and indicate that regional vegetation cycles closely accompany deglacial pulses of global extent.

The stratigraphy of core KET 8003 (38°49.2' N, 14°29.5' E, 1,900 m water depth, 1,030 cm length), is based on the oxygen isotope variation in the foraminifer *Globigerina bulloides*, measured at 10-cm intervals⁴ (Fig. 1). Tests of calcite formed by marine organisms reflect the fluctuations of ^{18}O in sea water. Because ^{16}O evaporates more readily than ^{18}O , the $^{18}\text{O}/^{16}\text{O}$ ratio is low in continental ice sheets and high in the glacial ocean. Conversely, during interglacial periods, the oceans are depleted in ^{18}O . The basis of this core lies within the last Interglacial, that is isotope stage 5e at 125 kyr BP. The chronology is based on six interbedded ash layers that have been correlated by their chemical composition with lava flows in southern Italy^{4,5} that were dated by ^{14}C and K/Ar analysis to be between 55 and 12 kyr old, and on the initiation of the deglaciation after the last glacial maximum (Termination I), globally seen at 15 kyr BP⁶. Because the sample set for pollen was taken subsequently, it has a 2-cm offset from the isotope set. Pollen has not been recovered below 580 cm.

Southern Italy was also the source area for pollen in core KET 8003: both volcanic ash and pollen are airborne. We have identified 128 pollen taxa and excluded *Pinus* from the pollen sum. The large production and strong resistance of *Pinus* pollen to aerobic bacterial degradation, probably linked to its high sporopollenin content⁷, often result in over-representation in marine cores⁸. Here we describe the variation of the pollen percentages of *Quercus*, *Abies*, *Populus*, *Poaceae*, *Artemisia*, *Chenopodiaceae* and *Ephedra*. A heavy pollen producer, *Quercus* (oak) is the only deciduous tree whose pollen is continuously present in the core. Oak pollen percentage in core KET 8003 shows no correlation with (1) the pollen sum, (2) total pollen concentration per gram of sediment and *Pinus* pollen abundance. These two observations indicate respectively that the percentage is not an artefact of the analysis procedure nor is it controlled by differential pollen preservation before and during burial on the sea floor.

The $\delta^{18}\text{O}$ and oak pollen abundance series of core KET 8003 co-vary in phase with each other. The peaks in oak pollen abundance usually just precede those of $\delta^{18}\text{O}$ depletion. This, however, is an artefact of the sampling offset and does not preclude the existence of coincident peaks. The partitioning of the records into two types of intervals, one of $\delta^{18}\text{O}$ decrease/oak pollen increase (climate improvement) and the other with the reverse relationship (climate deterioration) yields a high correlation between the two data sets during the former intervals ($r=0.84$), and a lower one during the latter ($r=0.55$).

During isotope stage 3, characterized by generally high $\delta^{18}\text{O}$ values which denote globally glacial conditions between 69 and 24 kyr BP, there are four closely matched events of $\delta^{18}\text{O}$ depletions and oak pollen peaks between 580 and 240 cm which can be seen as short-lived deglacial pulses. Each of the two oldest peaks in oak pollen abundance, dated before 50 kyr BP, is set between two other abundance peaks: an earlier one of *Poaceae* and a later (and larger) one of *Artemisia*. This repetitive succession amounts to a pattern also seen later, during the first half of the deglaciation after 15 kyr BP. Whereas the abundance of oak pollen decreases above 472 cm, that of *Abies* pollen increases strongly between 450 cm and 400 cm, and *Hedera* and *Ilex* pollen appear sporadically. We suggest a correlation with the peaks of *Abies* pollen seen between 2,200 and 2,045 cm in the record of Lake Monticchio in southern Italy⁹. After 33 kyr BP, *Quercus* and *Artemisia* pollen abundances become minimal and *Poaceae* is the prevailing herbaceous pollen taxon.

During the glacial maximum (240–190 cm), corresponding to part of isotope stage 2 (24–11 kyr BP), the average abundance of oak pollen is at its lowest. Until mid-stage 2, the dominant

non-arboreal pollen remains Poaceae, and the abundances of *Artemisia*, Chenopodiaceae and *Ephedra* pollen are minimal.

The global decrease in ice volume starts at 190 cm (thus dated 15 kyr BP); this isotope signal precedes the expansion of oak pollen abundance which begins at 14 kyr BP, as in Tenaghi Philippon in Greece¹⁰. During this 1,000-yr timespan, the pollen abundance of aspen (*Populus*) is highest (up to 6%), and that of Poaceae is still moderately high. Simultaneously with the $\delta^{18}\text{O}$ reversal seen soon after 12.97 ± 0.18 kyr BP, oak pollen abundance decreases abruptly by one-third. Finally, the oak pollen abundance increase and deglaciation resume synchronously at the beginning of the Holocene.

As the oak pollen abundance rises after the glacial maximum, the high pollen abundance of Poaceae is progressively replaced by that of *Artemisia*, Chenopodiaceae and *Ephedra*. The abundances of these three taxa briefly decline at 122 cm with the first peak of oak pollen, and the largest $\delta^{18}\text{O}$ decrease, and culminate at 112 cm during the oak pollen minimum and $\delta^{18}\text{O}$ reversal. The isotope sequence points to the two stages of Termination I, Ia and Ib, separated by an abrupt and brief glacial reversal^{6,8}. The Tyrrhenian Sea record establishes stratigraphically that the glacial regrowth seen after 12.97 ± 0.18 kyr BP is synchronous with a decline in arboreal vegetation. The timing and character of these events, documented at a 1–6-cm sampling interval, support their identification with the Younger Dryas, which was a 600–900-yr period of cold climate between 11 and 10 kyr BP^{6,11}. The peak in oak pollen abundance at 122 cm is therefore corre-

lated with the Bölling–Alleröd, a warm climate period preceding the Younger Dryas. In the upper 90 cm of the core, trilete spores, mostly *Pteridium* (not represented in Fig. 1), become very abundant. This increase indicates the opening of the vegetation cover¹², possibly by humans.

Modern analogues for the pollen spectra in core KET 8003 are located in southeastern Europe and Middle Asia, 9–16° further north than the core site. Climate data in these regions support the hypothesis that the source area vegetation for this record responded primarily to moisture availability^{10,13,14}. The dominant type of oak pollen in the core is *Quercus robur pubescens*^{15,16}. One of the three most widespread species in Italy¹⁷, *Q. pubescens* dominates the temperate cool forests of the sub-humid and humid Mediterranean altitudinal zones, where precipitation is 600–1,200 mm yr⁻¹, below the *Fagus* and *Abies* forests of the cloud belt¹⁸. In the pollen assemblage of core KET 8003 where Poaceae or *Artemisia*–Chenopodiaceae pollen are prevalent, increases of oak pollen abundance can be accounted for by a precipitation increase or decrease in summer aridity in the 200–1,300 m elevation range. The six other reported taxa characterize today's vegetation formations of southeastern Europe and Middle Asia¹⁹: in the lowlands, *Populus* from the forest-steppe; Poaceae from the grass steppe; Chenopodiaceae, *Artemisia* and *Ephedra* from the continental semi-desert; and

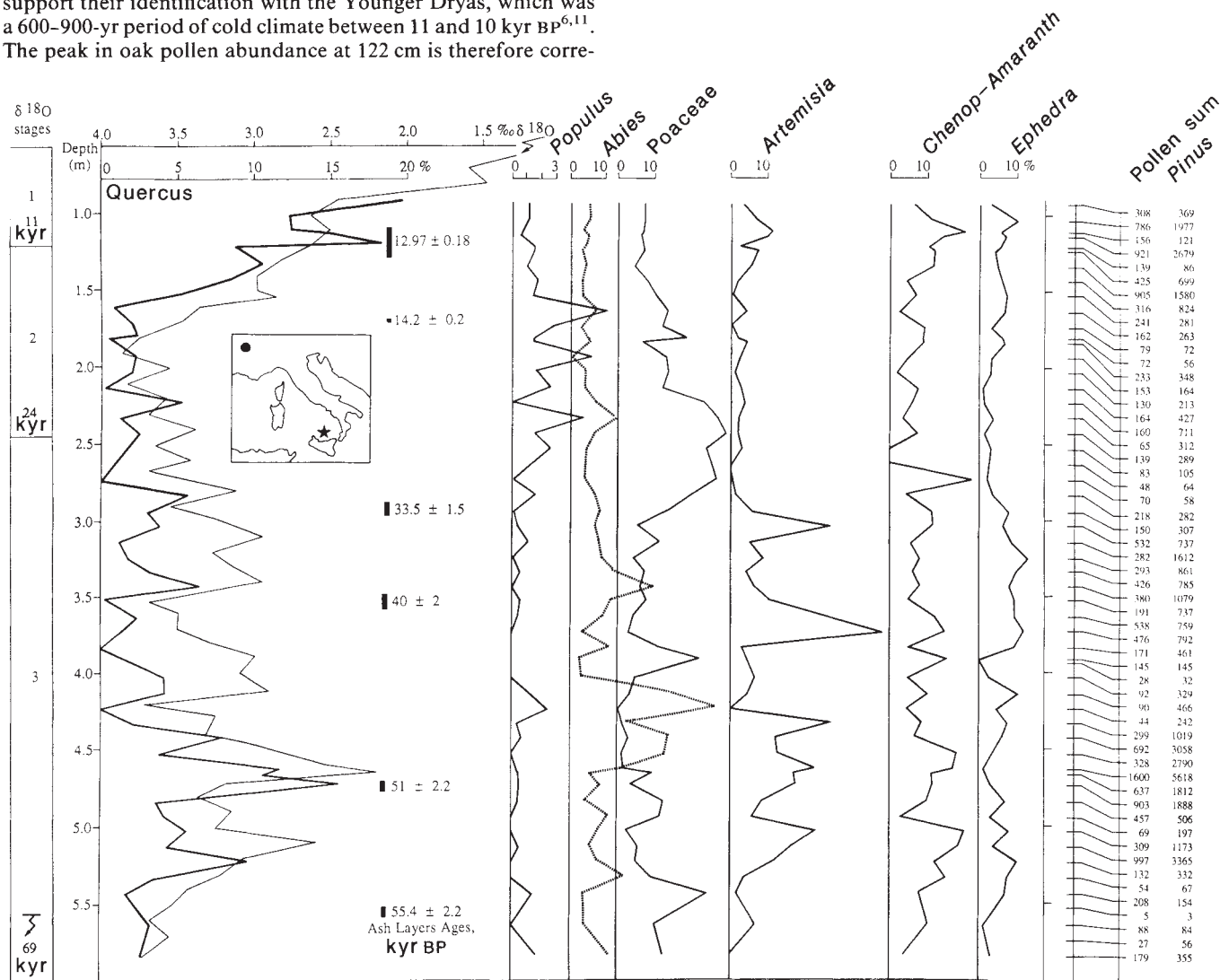


FIG. 1 Variation of $\delta^{18}\text{O}$ (‰) in *Globigerina bulloides*⁴ and pollen abundance (%) as a function of depth in core KET 8003 (80 km N of Sicily and 100 km W of Italy) from 55 to 9 kyr BP. *Pinus* is excluded from the pollen sum. Insert map: black star, location of core KET 8003; black circle, location of Les

Echets. Scales for $\delta^{18}\text{O}$ and oak pollen adjusted to allow easy visual comparison. Radiometric dates measured on lavas in southern Italy and peripheral islands^{4,5}. The six ash layers in the core are chemically correlated to these dated lavas.

in the uplands, *Abies* from the cool humid montane forest. A dense, humid grass steppe is seen today from Hungary through the Ukraine (annual precipitation 350–455 mm) along 45–50° N, to northern Kazakhstan at 50–55° N (ref. 19). With a thick snow cover providing ample spring moisture, the Ukrainian climate is partly oceanic, and with a precipitation peak in summer, it is partly continental. The late summer drought eliminates deciduous trees, leaving the steppe a residual vegetation formation. The lowland semi-desert seen today in Middle Asia along 50° N is more arid (80–280 mm precipitation) and more continental (21–24 °C in July, –12 to –16 °C in January) than the Ukrainian grass steppe. Forests of deciduous trees below *Abies* forests today grow on the northern and western slopes of the Tian-Shan mountains above the desert lowland that lies south of the semi-desert.

We interpret the vegetation that produced the pollen record from stage 3 as repeatedly shifting, in the lowlands, from semi-desert to steppe, with a predominance of the former under a highly continental climate. The succession of abundance peaks for oak pollen indicates that pulses of moisture and temperature increase occurred in the uplands, followed by episodes of still wet but colder climate, signalled by the high abundance of *Abies* pollen. The rapid rise in oak pollen abundance after each glacial interval indicates the ability of that tree to take advantage of the earliest moisture increase in a continental climate. This can be accounted for by the 'hardening' process: the survival of dormant buds is improved by low temperature (–25 °C and even –30 °C) when it occurs in January and February¹⁹.

The scarcity of oak pollen during the last glacial maximum indicates that pollen production decreased and/or that source areas were very restricted. Isolated refuges of deciduous temperate forest have been hypothesized in the mountains of southern Europe, low enough to avoid extreme autumn cold, fatal to the dormant buds, and high enough to provide adequate summer moisture by adiabatic cooling of ascending air^{13,14,20,21}. The replacement of the present Mediterranean vegetation of southern Italy at sea level by a grass steppe of Ukrainian type would result from a temperature decrease of 1–4 °C in July, to 24–21 °C, and of 10–24 °C in January, to 0 to –14 °C, and a 440–540-mm decrease in annual precipitation, to 450–350 mm. Provided that precipitation would be ≥ 500 –600 mm, then a refuge zone for deciduous trees could exist up to 600 m, where temperature would be 17.4–20.4 °C in July, –3.6 to –17.6 °C in January (with a 0.6 °C per 100-m pseudoadiabatic lapse rate²²). These temperatures compare well with those at the northern and north-east limits of the deciduous oak distribution in eastern Europe¹⁹. Thus it might be that the lower temperatures during glacial times would not have restricted or eliminated the oak, whereas moisture decrease could have, as also suggested by estimates of glacial climate in Greece¹⁴.

A modern analogue for the vegetation that produced maximum *Populus* pollen abundance at the onset of deglaciation in core KET 8003 is the aspen trees, which in the eastern Ukrainian steppe surround each small ground depression occupied by a rain-fed round lake (called a 'pod')¹⁹. Although higher than in the treeless steppe, precipitation (~480 mm) is still not sufficient to develop a drainage system.

During the early part of deglaciation, ~14 kyr BP while the upland oak population increased in the Tyrrhenian basin, the lowland oceanic grass steppe of the glacial maximum progressively gave way to a continental semi-desert, most widespread during the Younger Dryas.

Reconstruction of the precipitation variation in Les Echets in south-east France²³ (Fig. 2) can be compared with our oak pollen per cent variation. In the two time-series, during stage 3, moderately high values are seen from 60 to 43 kyr BP, and low values at ~40 kyr BP, increasing during the last part of stage 3. Stage 2 is drier and oak values are minimal; precipitation starts to increase at 15 kyr BP, as to the deciduous trees.

The climatic significance of this isotope record is complex.

The 1.3‰ isotope effect of the global ice volume²⁴ accounts for 54% of the 2.40‰ $\delta^{18}\text{O}$ range between glacial maximum and the core top. The 1.1‰ remainder is due to regional sea surface temperature and salinity variations. As the isotope and oak pollen variations have comparable amplitudes in the strong deglacial pulse of early stage 3 from 560–542 to 472–466 cm depth ($\Delta(\delta^{18}\text{O}) = 1.48\text{‰}$, Δ oak pollen = 14%) and in the transition from glacial maximum at 190 cm to Alleröd at 110–120 cm (Termination Ia) ($\Delta\delta^{18}\text{O} = 1.37\text{‰}$, Δ oak pollen = 16%), it is conceivable that the isotope shift of stage 3 could record an ice-volume decrease comparable with that of the first phase of deglaciation. This implication, which requires that the range of salinity variation is also similar, is based on the present link between the formation or rejuvenation of atmospheric depressions, generating the regional precipitation to which the oak population is sensitive, and the high Tyrrhenian sea surface temperature^{25,26}.

Comparison of the course of events after the deglacial pulses of stage 3 and Termination Ia shows that the reversal to glacial conditions in the isotopic composition of the sea and in the vegetation of southern Italy was complete after 49 kyr BP, and partial after 11 kyr BP during the Younger Dryas. We hypothesize that this difference might have some link with the observations that (1) the July insolation peak at 56 kyr BP was lower than that at 11 kyr BP (at 60° N, respectively 7 and 10% higher than present²⁷) and (2) the rate of insolation increase for July was smaller at the transition from full-glacial stage 4 to stage 3 than at that from full-glacial stage 2 to stage 1 (at 60° N, respectively 7.2 cal cm⁻² per 1,000 yr between 70 and 56 kyr BP, and 8.5 cal cm⁻² per 1,000 yr between 22 and 10 kyr BP). The Termination II deglaciation between 1,350 and 1,250 kyr BP does not, to present knowledge, record a reversal similar to the Younger

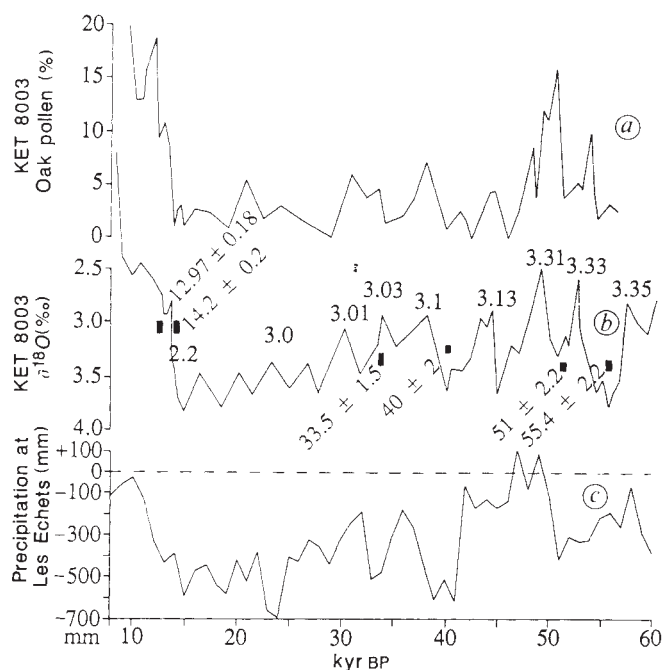


FIG. 2 a, Time series of oak pollen in core KET 8003. b, Time series of $\delta^{18}\text{O}$ in core KET 8003. Chronostratigraphy based on radiometrically dated tephras (black rectangles) and the initiation of deglaciation at 15 kyr BP, isotope event 2.2. Event 3.0, the stage 2/3 boundary, is set at 24 kyr BP. Isotope events after ref. 5. c, Reconstruction of variation in annual total precipitation expressed as deviation from the modern value (800 mm) in Les Echets, southeastern France²³. Error bars omitted for simplification. On comparing this with the KET 8003 time series, a lack of radiometric ages in Les Echets before 25 kyr BP should be noted. The major precipitation peak in Les Echets at 49–46 kyr is synchronous with the peak in *Abies* pollen abundance in KET 8003.

Dryas²⁸. The link of Termination II with a higher insolation peak and more rapid pace of insolation increase is consistent with the steering of the climate system.

Broecker hypothesized that it is the shut down of the Atlantic Ocean conveyor belt that brought about the cooling trend in the climate jumps after 11 kyr BP and during stage 3 (refs. 29 and 30). If this is the case, the heat that is no longer transported from low to high latitudes by ocean surface currents would accumulate in the tropical ocean. There it would increase the convective activity of the Intertropical Convergence Zone, which would reactivate the atmospheric Hadley cell^{31,32} and restore the poleward oceanic heat transport. Thus the conveyor belt closing-off is a self-limiting process which can only be

temporary.

Inclusion in rapidly accumulating marine sediments of the pollen originating in all elevations in the rimlands of the Tyrhenian Sea is a process that provides an integrated picture of the vegetation through its rapid changes. The complex pattern of these changes documents the palaeoclimate of southern Europe during a timespan that is here particularly well dated. The covariation of the oak pollen abundance with the isotopic composition of sea water, which is shaped in part by the global volume of the ice sheets, testifies to the coherence of the fluctuations of the climate system at the millennial timescale in its regional and global expressions. □

Received 21 February; accepted 9 October 1989.

- Heusser, L. & Shackleton, N. J. *Science* **204**, 837–839 (1979).
- Van Campo, E., Duplessy, J. C. & Rossignol-Strick, M. *Nature* **269**, 56–59 (1983).
- Turon, J. L. *Nature* **309**, 673–676 (1984).
- Paterne, M., Guichard, F., Labeyrie, J., Gillot, P. Y. & Duplessy, J. C. *Mar. Geol.* **72**, 259–285 (1986).
- Paterne, M., Guichard, F. & Labeyrie, J. *J. Volcan. geotherm. Res.* **34**, 153–172 (1988).
- Bard, E. *et al.* *Nature* **328**, 791–794 (1987).
- Shaw, G. in *Sporopollenin* (eds Brooks, J., Grant, P. R., Muir, M. D. & Van Gijzel, P.) 305–348 (Academic, London, 1971).
- Duplessy, J. C., Delibrias, G., Turon, J. L., Pujol, C. & Duprat, J. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **35**, 121–144 (1981).
- Watts, W. A. *J. geol. Soc. Lond.* **142**, 491–499 (1985).
- Wijmstra, T. A. *Acta bot. neerl.* **18**, 511–527 (1969).
- Cwynar, L. C. & Watts, W. A. *Quat. Res.* **31**, 377–380 (1989).
- Behre, K. E. *Pollen Spores* **23**, 225–245 (1981).
- Smit, A. & Wijmstra, T. A. *Acta bot. neerl.* **19**, 867–876 (1970).
- Sabatier, R. & Van Campo, M. *Bull. Soc. bot. Fr.* **131**, 85–96 (1984).
- Planchais, N. *Pollen Spores* **4**, 87–93 (1962).
- Smit, A. *Acta bot. neerl.* **22**, 655–665 (1973).
- Bonin, G. & Gamisans, J. *Doc. Phytos.* **19–20**, 73–88 (1976).

- Quézel, P. *UNESCO Note techn. MAB* **2**, 9–33 (1976).
- Walter, H. *Die Vegetation Osteuropas, Nord- und Zentralasiens* (Fischer, Stuttgart, 1974).
- Beug, H. J. *Ber. dt. bot. Ges.* **80**, 682–689 (1967).
- Van Campo, M. in *Études françaises sur le Quaternaire* (Suppl. *Bull. AFEQ*, Paris) 104–111 (1969).
- Holton, J. R. *An Introduction to Dynamic Meteorology* (Academic, New York, 1979).
- Guiot, J., Pons, A., de Beaulieu, J. L. & Reille, M. *Nature* **338**, 309–313 (1989).
- Ruddiman, W. F. in *North American and Adjacent Oceans During the Last Deglaciation* Vol. K3 (eds Ruddiman, W. F. & Wright, H. E. Jr) 463–478 (Geol. Soc. Am., Boulder, 1987).
- Trewartha, G. T. *The Earth's Problem Climates* (University of Wisconsin Press, Madison, 1961).
- Gat, J. R. & Carmi, I. *J. Geophys. Res.* **75**, 3039–3048, 1970.
- Berger, A. *Quat. Res.* **9**, 139–167 (1978).
- Wollard, G. M. *Quat. Res.* **9**, 1–21 (1978).
- Broecker, W. S., Peteet, D. M. & Rind, D. *Nature* **315**, 21–26 (1985).
- Broecker, W. S. *et al.* *Paleoceanography* **3**, 1–19 (1988).
- Bjerknes, J. *Tellus* **18**, 820–829 (1966).
- Rowntree, P. R. *Q. Jl. R. met. Soc.* **98**, 290–321 (1972).

ACKNOWLEDGEMENTS. This work was supported by CNRS, PNEDC and EEC (Climate programme) and Cruise ETNA-80 of the French R/V Noroit, by CNEXO. We thank M. Paterne for the isotope and radiometric data, J. Guiot for the precipitation estimate and L. Mayer for help with the manuscript.

Seismic reflection profiles across deep continental crust exposed in the Kapuskasing uplift structure

J. A. Percival*, A. G. Green*, B. Milkereit*,
F. A. Cook†, W. Geis† & G. F. West‡

* Geological Survey of Canada, 601 Booth Street, Ottawa, Ontario, Canada, K1A 0E8

† Department of Geology and Geophysics, University of Calgary, Calgary, Alberta, Canada, T2N 1N4

‡ Department of Physics, University of Toronto, Toronto, Ontario, Canada, M5S 1A7

SEISMIC imaging of Precambrian cratons holds the promise of deep structural mapping and interpretation of fundamental crustal construction processes. Deep structures may be identified by tracing reflections to the surface¹, particularly in exposed crustal cross-sections² elevated along deeply penetrating faults, but many such cross-sections have structural orientations too steep for successful reflection imaging (like the Ivrea zone³) or occur in plate-boundary settings not representative of continental interiors. By contrast, the Archaean greenstone terrain exposed in the Kapuskasing uplift of Canada^{4,5} has gently dipping structures, and an intracratonic setting within the Superior Province, thus providing an opportunity to examine the third dimension of continental crust down to 25–30 km palaeodepth⁶. Here we present new seismic reflection data which, although generally supportive of structural models based on surface geology⁴, gravity modelling⁵ and seismic-refraction studies^{7,8}, also indicate that the Kapuskasing structure is a relatively thin thrust sheet containing seismically reflective sequences of high-grade rock which may be analogues of seismically reflective *in situ* lower crust.

Superior Province, the largest preserved Archaean craton, is

well known for its greenstone belts, hosting major gold and base-metal deposits. The east-trending Abitibi–Wawa belt (Fig. 1) of low-grade metavolcanic and granitoid rocks is transected by the north-east-striking Kapuskasing uplift (KU), distinguished by high-grade metamorphic rocks producing positive gravity and aeromagnetic anomalies⁹. The KU is interpreted as a slab of deep crust^{4,10,11}, thrust upward along the west-dipping Ivanhoe Lake fault zone (ILFZ). A 120-km transect across the KU from Lake Superior to the ILFZ crosses greenstones with 2–3 kbar metamorphic palaeopressures¹², amphibolite-facies (5–6 kbar) tonalite of the Wawa gneiss terrane, a density¹³–velocity¹⁴ boundary inferred to be a Conrad-type discontinuity, and heterogeneous, layered, high-grade (7–9 kbar) gneisses of predominantly magmatic origin in the Kapuskasing zone⁶. Increases in metamorphic grade and pressure across the KU indicate an easterly deepening erosion level, from regional background values of 6–9 km in the greenstones to 25–30 km near the ILFZ. With increasing palaeodepth, decreasing U–Pb, Ar–Ar and Rb–Sr ages show that rocks from deeper levels remained hot for longer and cooled more slowly¹⁵. The crust attained most of its 40±5-km thickness in the interval 2.75–2.68 Gyr, through early arc volcanism and plutonism, followed by shortening, metamorphism and renewed arc magmatism⁵. Seismic refraction results map a zone of high velocity (6.6 km s⁻¹) extending from near the surface in the Kapuskasing zone to normal depths of ~20 km under the greenstones to the west^{7,8}. The crust thickens from regional values of 35–44 km to >50 km beneath the Kapuskasing zone.

Since 1984 LITHOPROBE has supported a wide variety of geological and geophysical studies aimed at defining crustal-scale geometry and evolution through the Kapuskasing 'window'. A key component of these studies is the 1987–88 seismic reflection survey; three lines, labelled 1, 2–3–4 and 6, crossed the southern KU (Figs 1 and 2). Line 1 extended south to north from the Abitibi belt into the Kapuskasing zone. Composite line 2–3–4, comprising three conventional segments and two under-shoots regions containing reflection points but having no surface reflection sources or receivers, crossed the Kapuskasing zone

including the Shawmere anorthosite complex parallel to the regional east-west structural grain. In addition to the relatively conventional regional survey, line 2 was resurveyed using high-resolution parameters, which yielded significantly more detail about the shallow structure (compare Fig. 3a and b). Finally, line 6 sampled the southern end of the KU.

The stacked seismic sections (Figs 2, 3) show many coherent events although the image quality varies along the profiles. Upper crustal reflectivity is generally high along all lines, but the reflection character varies with the geological terrane: Portions of the seismic lines that cross tonalitic gneiss and granite of the Abitibi belt yield prominent reflections, dipping west and north at $>30^\circ$, from near the surface to at least 13 s beneath the surface exposure of the KU. The uppermost zone of strong reflectivity on lines 1 and 6 (H on Fig. 2a, c) projects 10 km beyond the south end of the lines into outcrops of felsic gneiss and 'homogeneous' granite.

Reflectivity of the KU granulites and the seismic image of the ILFZ are best represented on line 2-3-4. Within the Kapuskasing zone, shallow (0.2-2 s) 'E' reflections (Fig. 3b) dip $\sim 15^\circ$ W along the profile and can be traced short distances to surface exposures of mafic, tonalitic and anorthositic rocks interlayered on a 10-30-m scale¹⁶. These sill-like intrusive bodies have seismic velocities (7.23, 6.55, 7.26 km s⁻¹ respectively)¹⁷ and dimensions appropriate to generate significant reflections although direct correlation of reflections and lithology would require short drill holes¹⁸. The basal reflection of the E package (I', Fig. 2b) locally truncates events above and below it, suggesting hanging-wall and footwall cutoffs by a splay of the ILFZ¹⁹. A pair of gently west-dipping reflections (A, B in Fig. 3) steepen upward from 2-3 s. Associated structures can be followed along distinct reflections or truncations of events to near the surface trace of the ILFZ (Figs 2b, 3), using geometry based on migrated sections. Here the reflections may correspond to zones of cataclase and mylonite up to 70 m thick¹¹. To the west, reflections A and B merge with the base of the E package to become a diffuse zone extending to ~ 4 s at the western end of the composite line. The ILFZ is not represented by distinct reflections on the other lines, but is constrained by non-offset and non-truncation of reflections to dip $\sim 15^\circ$ N along line 1 and $\sim 15^\circ$ NW along line 6. The inferred extension of the ILFZ occurs at a common depth of ~ 3.8 s at the intersection of the lines and possibly continues to ~ 5 s at the north end of line 1 (I, Fig. 2a).

In the central and western parts of the survey area, the Kapuskasing zone and Wawa gneiss terrane are represented down to 6-8 s by discontinuous, rolling sub-horizontal reflections up to 5 km long (Fig. 2). Again, shallow reflections can be related to rock exposures. A zone of sub-horizontal reflectivity at 0.3-3 s at the northern end of line 1 (F, Fig. 2a) is consonant with gently undulating surface structure in the overlying high-grade paragneiss; concave-down reflective zone G occurs beneath a structural dome of tonalitic gneiss. Shallow reflections (J) from the thin wedge of hanging-wall rocks at the north-west end of line 6 are sub-horizontal or gently west dipping, corresponding to surface structural geometry²⁰.

The depth to the base of strong reflectivity varies along each line and between lines. Beneath line 2-3-4 it occurs at 13 s in the east, at 6 s in the east central region, and undulates between 11 and 13 s elsewhere. Reflections are abundant to at least ~ 13 s beneath the Abitibi belt on lines 1 and 6, but decrease markedly at 8-9 s to the north. The variation in depth extent of reflectivity does not seem to correlate with crustal thickness, estimated at $14-16 \pm 0.6$ s over the area of the reflection survey (Fig. 1)^{7,8}. Although some of the change in reflectivity is attributable to varying surface conditions, it is possible that a deep-crustal deformation zone, characterized by disrupted structures and weakened reflections, characterizes the crustal bulge beneath the KU.

The reflection sections have in common: (1) shallow (<2.0 s) sub-horizontal or west-dipping reflections that can be projected short distances to surface structure in both the granulite-facies Kapuskasing zone and amphibolite-facies Wawa gneiss terrane; (2) images of the ILFZ, a structure dipping 16° north-north-west that can be traced to 4-5 s; (3) prominent west or north-dipping reflectors beneath the Abitibi subprovince; (4) reflections to ~ 13 s in the Abitibi crust by contrast with absent or less coherent reflections below 6-8 s beneath the KU; and (5) no discrete reflecting zone or marked reduction in reflectivity corresponding to the refraction-determined Moho.

Based on the reflection image of the ILFZ, the KU is a much thinner feature than inferred from earlier gravity modelling⁴. Traced to 4 s on line 2-3-4 and 5 s on line 1, the fault probably continues to dip northwesterly to a depth of ~ 20 km (6.6 s) based on refraction data from further west⁷. These constraints allow us to estimate the horizontal shortening required to elevate the granulites to the surface. Their depth of origin is ~ 20 km,

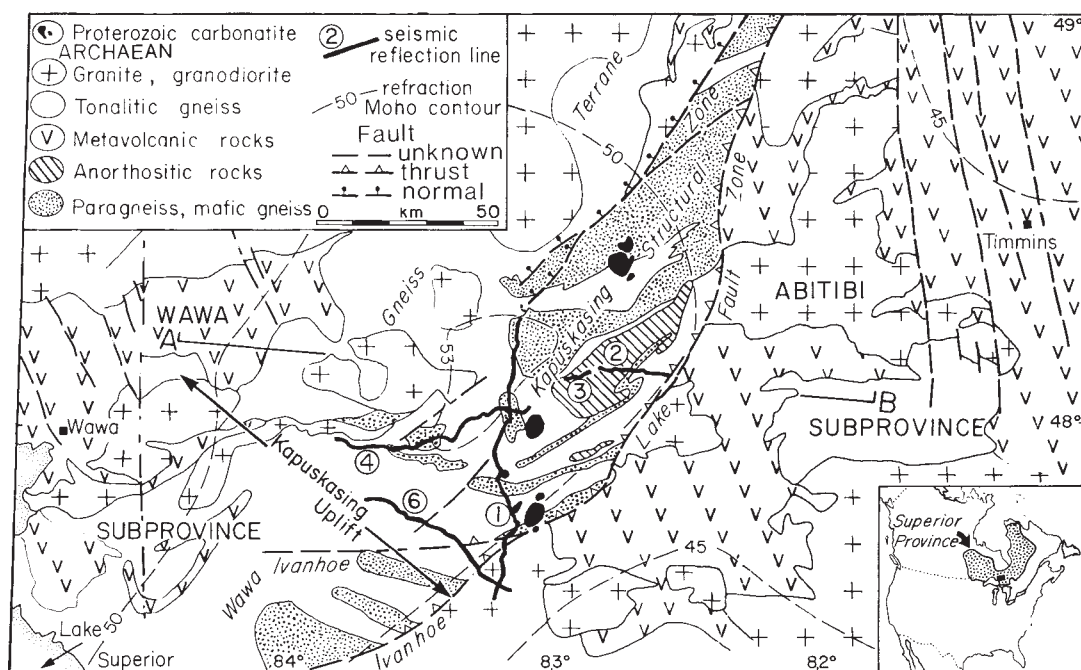


FIG. 1 Simplified regional geological map with locations of seismic-reflection lines and Moho contours⁸ in km.

based on the palaeopressure difference of ~ 6 kbar between the low-grade rocks near Lake Superior and the Kapuskasing granulites, and on refraction data, linking the 6.6 km s^{-1} velocities of the Kapuskasing zone to normal depths for this velocity (20 km) near Lake Superior. The shortening estimate then depends on assumptions concerning fault geometry north-west of the region surveyed. If the fault dips consistently at 16° to 20-km depth, approximately 70 km of shortening is required, whereas the presence of a ramp near the north-west edge of our survey¹⁹ would lessen this to a minimum value of 55 km.

Similar magnitudes of overthrusting have been documented in many orogenic belts²¹, where décollements must extend hundreds of kilometres to plate boundaries to balance shortening on a lithospheric scale. Although the crustal geometry is rela-

tively well constrained, the lateral continuation of the ILFZ is undetermined and therefore the mechanism of crustal deformation that caused the KU is not fully resolved. Structural models that account for the shortening and associated >10 -km increase in crustal thickness (Fig. 4) include: (1) the detachment roots at an ancient plate boundary to the north-west¹⁹; (2) convergence from the south-east caused underthrusting of the Abitibi rocks beneath the KU (tectonic delamination²²); and (3) brittle shortening of the upper crust was balanced by ductile flow in the lower crust (as in the Zagros collision zone²³).

Although the seismic data provide no direct evidence on the age of thrusting, the inferred 55–70 km of shortening has implications for the timing. The arcuate 2.45-Gyr (ref. 24) Matachewan–Hearst dyke swarm crosses the southern KU with

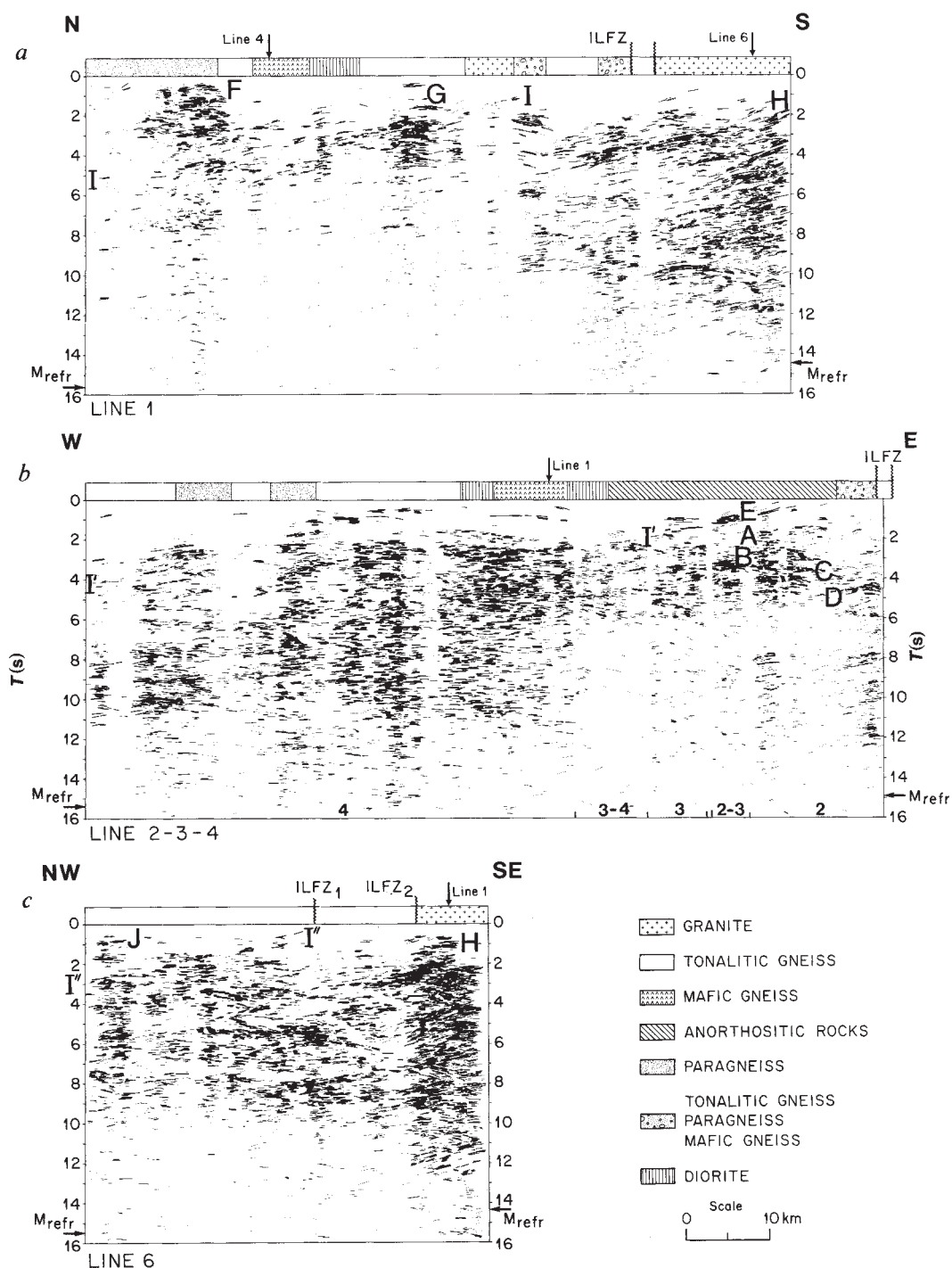


FIG. 2 Line drawings of *a*, line 1 *b*, line 2-3-4 and *c*, line 6 based on unmigrated 60-fold data; an automated procedure produced similar results. Parts of 2-3-4 indicated by 2-3 and 3-4 are undershoots. Location of the lines is shown in Fig. 1. ILFZ₁ and ILFZ₂ identify two splays of the Ivanhoe Lake fault zone. A portion of the composite line 2-3-4 is shown in Fig. 3. Vertical scale is travel time in seconds; approximate depth in kilometres may be obtained by multiplying travel time by three. The sections are plotted at 1:1 for a velocity of 6 km s^{-1} . Annotation refers to features discussed in the text.

FIG. 3 Comparison of regional and high-resolution LITHOPROBE stacked seismic reflection data collected along the eastern segment of composite line 2-3-4. Location of the line is shown in Fig. 1 and the geological legend in Fig. 2. Geology is taken from more detailed maps than shown in Fig. 1. The regional data were collected with the following parameters: 50-m receiver spacing, 100-m source spacing, 240 channels, 60-fold coverage, 4 large vibrators, 12–52 Hz sweeps, 16-s correlated record length, 4-ms sample interval. The high-resolution data were collected with the following parameters: 20-m receiver spacing, 20-m source spacing, 120 channels, 60-fold coverage, 2 large vibrators, 20–130 Hz sweeps, 8-s correlated record length, 2-ms sample interval. Plotting parameters as in Fig. 2. Annotation refers to features discussed in the text.

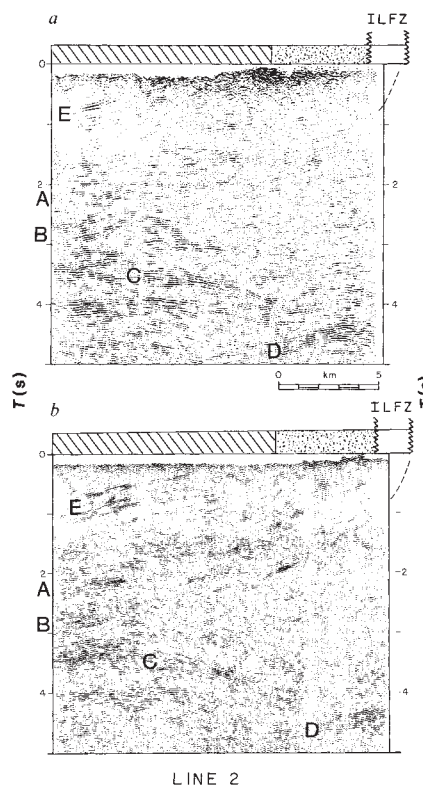
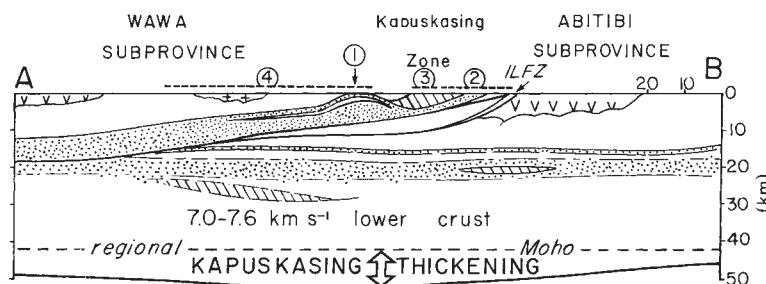


FIG. 4 Summary cross-section through the Kapuskasing uplift (legend as in Fig. 1), showing 70 km of horizontal shortening by thrusting and ductile thickening, respectively above and below a décollement at ~20 km depth. 1 to 4 (circled) show approximate locations of the seismic reflection profiles relative to the sketch. Alternative mechanisms that account for the geometry are discussed in the text.



only minor offsets²⁵. Unless the transport direction coincidentally followed the arcuate trend of the dykes, this suggests that major movement preceded 2.45 Gyr, supporting some palaeomagnetic²⁶ and Ar-Ar age²⁷ observations and contrasting with inferences^{15,28} of principal uplift at ~2 Gyr.

To generally apply the model of crustal evolution developed through the three-dimensional Kapuskasing exposure, we must compare the KU seismic signature to that of other regions. The general character of the Kapuskasing profiles, with abundant reflections in the upper 8 s, diminishing downwards to ~13 s, some distance above the refraction-determined Moho, is similar to other Precambrian areas²⁹ but differs from the pattern observed in some younger terranes of transparent upper and reflective lower crust³⁰. Several explanations for these crustal-scale differences are possible: (1) reflectivity is enhanced by some parameter that changes with time (slowly released metamorphic fluids, for example); (2) the Kapuskasing area has a deep, nonreflective component not found in younger crusts (such as a massive underplate); (3) information from young crusts is biased towards extended terranes that have a characteristic structure; or (4) there are fundamental differences in the structure and evolution of Archaean and younger crusts. Nevertheless, for the kilometre-scale, laminated, deep-crustal reflections that make up the regional patterns, the reflective, sub-horizontally-layered, high-grade rock sequences at Kapuskasing, with velocities in the 6.6 km s⁻¹ range, provide a viable analogue. Arc-building processes involving injection of sub-horizontal sheets of mafic and felsic magma, with attendant differentiation,

metamorphism and assimilation, may therefore have been widespread during construction of the foundations of the continents. □

Received 26 June; accepted 16 October 1989.

1. Oliver, J. E., Brown, L. D. & Cook, F. A. *J. geophys. Res.* **88**, 3329–3347 (1983).
2. Fountain, D. M. & Salisbury, M. H. *Earth planet. Sci. Lett.* **5**, 263–277 (1981).
3. Bayer, R. *et al. C.R. heb. Séanc. Acad. Sci. Paris Sér. II*, **305**, 1461–1470 (1987).
4. Percival, J. A. & Card, K. D. *Geology* **11**, 323–326 (1983).
5. Percival, J. A. & Card, K. D. in *Evolution of Archean Supracrustal Sequences*, *Geol. Ass. Can. Spec. Pap.* **28**, 179–192 (1985).
6. Percival, J. A. *Am. Miner.* **68**, 667–686 (1983).
7. Boland, A. *et al. Nature* **335**, 711–713 (1988).
8. Boland, A. & Ellis, R. M. *J. geophys. Res.* **94**, 7189–7204 (1989).
9. Percival, J. A. & McGrath, P. H. *Tectonics* **5**, 553–572 (1986).
10. Cook, F. A. *Geology* **13**, 368–371 (1985).
11. Bursnell, J. T. in *Geol. Surv. Pap. Can.* **89-1C**, 405–411 (1989).
12. Studemeister, P. A. *Can. J. Earth Sci.* **20**, 1409–1420 (1983).
13. Percival, J. A. *Am. geophys. Un., Geodyn. Ser.* **14**, 135–141 (1986).
14. Percival, J. A. & Fountain, D. M. in *Fluid Movements—Element Transport and the Composition of the Deep Crust* (ed. Bridgwater, D.) 51–60 (Kluwer, Dordrecht, 1989).
15. Percival, J. A., Parrish, R. R., Krogh, T. E. & Peterman, Z. E. in *Proc. Kapuskasing Structural Zone Transect Workshop* (ed. West, G. F.) 43–47 (University of Toronto, 1988).
16. Green, A. G. *et al. Tectonophysics* (in the press).
17. Fountain, D. M., Salisbury, M. H. & Percival, J. A. *J. geophys. Res.* (in the press).
18. Percival, J. A. *Geosci. Can.* **15**, 151–153 (1988).
19. Geis, W. T. *et al. Geology* (in the press).
20. Moser, D. E. in *Geol. Surv. Pap. Can.* **89-1C**, 215–224 (1989).
21. Cook, F. A. *et al. Geol. Soc. Am. Bull.* **92**, 738–748 (1981).
22. Meissner, R. *Terra Rev.* **1**, 17–28 (1989).
23. Snider, D. B. & Barazangi, M. *Tectonics* **5**, 361–373 (1986).
24. Heaman, L. D. *Geol. Ass. Can. Abstr.* **13**, A53 (1988).
25. West, G. F. *Geol. Ass. Can. Abstr.* **14**, A38 (1989).
26. Symons, D. T. A., Lewchuk, M. E. & Vandell, T. A. in *Proc. Kapuskasing Structural Zone Transect Workshop* (ed. West, G. F.) 205–210 (Univ. of Toronto, 1988).

27. Queen, M., Hanes, J. A. & Archibald, D. A. *Geol. Ass. Can. Abstr.* **14**, A123 (1989).
 28. Costanzo-Alvarez, V. & Dunlop, D. J. *J. geophys. Res.* **93**, 9126–9136 (1988).
 29. Gibbs, A. K. *Am. Geophys. Un., Geodyn. Ser.* **14**, 95–106 (1986).
 30. Matthews, D. H. & Cheadle, M. J. *Am. Geophys. Un., Geodyn. Ser.* **13**, 5–19 (1986).

ACKNOWLEDGEMENTS. We thank P. Hurley and R. Bottomley for technical assistance. Seismic reflection data were acquired and initially processed by Veritas Geophysical and Veritas Seismic under contract to LITHOPROBE. Funding to LITHOPROBE is provided by the Natural Sciences and Engineering Research Council of Canada and the Geological Survey of Canada.

The hafnium paradox and the role of garnet in the source of mid-ocean-ridge basalts

Vincent J. M. Salters*† & Stanley R. Hart*‡

* Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

MID-OCEAN-RIDGE basalts (MORBs) are thought to result from melting in the mantle at depths of less than 60 km, in the spinel stability field^{1–3}. MORBs have ¹⁷⁶Hf/¹⁷⁷Hf ratios indicating derivation from a mantle reservoir with a long-term Lu/Hf ratio greater than that of C1 chondrite meteorites, yet the measured Lu/Hf ratios in MORB are lower than in C1 chondrites: this is the 'hafnium paradox'⁴. Here we show that the Lu–Hf and Sm–Nd systematics of MORBs require garnet to be a residual phase in MORB melt genesis. This places the onset of melting beneath a mid-ocean ridge at depths greater than 80 km. A sequential melting model, in which melting starts in the garnet stability field and then continues at shallower levels, best explains the combined Nd and Hf isotope systematics, and is compatible with our present geophysical and geochemical knowledge of mid-ocean-ridge magmatism.

Patchett and co-workers^{5–7}, using hafnium isotopes to constrain geological processes, concluded that, in general, Hf isotopes in oceanic volcanics (MORB and OIB) are well correlated with Nd isotopes⁵. The Lu–Hf isotope system has therefore received only very limited attention in mantle studies. The good correlation between the Nd and Hf isotope systematics is believed to reflect similar behaviour of the parent/daughter ratios, Sm/Nd and Lu/Hf respectively, during mantle differentiation processes. We can use this similarity in behaviour and any deviation therefrom to constrain the MORB melting process.

MORBs are generally thought to form by decompression melting of a spinel lherzolite, followed by fractionation in a magma chamber before eruption. This model implies that the mantle residual to MORB genesis consists of olivine, orthopyroxene and probably some chromium diopside⁸. Estimates for the degree of melting of MORBs ranges from 1–25%^{8–10}. Most MORBs are believed to be generated at pressures of at least 8–10 kbar, as the melt is not in equilibrium with orthopyroxene^{3,11} below this pressure. In the anhydrous melting of spinel peridotite, the partition coefficients increase in the order Nd, Hf, Sm, Lu^{12,13}. Knowledge of the solid/melt distribution coefficients allows coupling of the isotopes, that is parent/daughter ratios in the mantle source can be related to their parent/daughter ratios in the melt. The coupling of the Nd and Hf systematics and the information they contain concerning the parent/daughter ratios of the source and melt have not been modelled before. MORBs form a fairly smooth trace-element pattern on incompatibility (spider) diagrams¹², indicating possible derivation from an anhydrous spinel lherzolite. Sm/Nd, Sm/Hf and Lu/Hf ratios in MORBs are not sensitive to shallow-level crystal fractionation effects, so they will reflect

the original source ratios as modified only by the melting process. The ¹⁷⁶Hf/¹⁷⁷Hf and ¹⁴³Nd/¹⁴⁴Nd ratios yield information on the long-term Lu/Hf and Sm/Nd ratios in the source.

Figure 1 shows the Lu/Hf and Sm/Nd ratios in MORBs and hypothetical isochrons for a MORB mantle source. All MORB values lie on the left of the 1–3 Gyr mantle isochrons, that is, the Lu/Hf and Sm/Nd ratios in MORB are too low for the observed ¹⁷⁶Hf/¹⁷⁷Hf and ¹⁴³Nd/¹⁴⁴Nd. This indicates that the process generating MORBs has produced significant fractionation (decrease) of the Lu/Hf and Sm/Nd ratios. This observation is the Hf paradox, in that the Lu/Hf in MORB is lower than C1 chondrite, whereas ¹⁷⁶Hf/¹⁷⁷Hf indicates a long-term Lu/Hf higher than C1⁴. As for Sm/Nd, the Lu/Hf ratio in C1 chondrite is also the best estimate for the bulk-Earth (BE) Lu/Hf ratio¹⁴.

To quantify these parent-daughter fractionations, we chose a 2-Gyr model age for the MORB source^{14,15}; a choice of 1.0 or 3.0 Gyr would not alter the arguments below. The difference between the mantle isochron value and the measured MORB value can be expressed as $\delta_{\text{Sm/Nd}} = ((\text{Sm/Nd})_{2\text{Gyr}} - (\text{Sm/Nd})_{\text{m}}) / (\text{Sm/Nd})_{2\text{Gyr}}$; where $(\text{Sm/Nd})_{2\text{Gyr}}$ is the calculated Sm/Nd based on the measured present-day ¹⁴³Nd/¹⁴⁴Nd of a basalt and $(\text{Sm/Nd})_{\text{m}}$ is the measured Sm/Nd in this basalt. $\delta_{\text{Lu/Hf}}$ is defined in the same way, and $\delta_{\text{Lu/Hf}}$ and $\delta_{\text{Sm/Nd}}$ for MORBs and OIBs are plotted in Fig. 2. A δ -value of zero means that the sample falls on a 2-Gyr isochron through BE (bulk Earth). The $\delta_{\text{Lu/Hf}}$ of MORB is by far larger than zero, which is the hafnium paradox. Some recent process must be responsible for these large δ -values.

The high $\delta_{\text{Sm/Nd}}$ in OIBs (that is, that the Sm/Nd ratios do not support the measured ¹⁴³Nd/¹⁴⁴Nd) was recognized early on¹⁶, and this phenomenon was shown to be more pronounced in the Lu–Hf system, extending to MORBs as well as OIBs⁵. The high $\delta_{\text{Sm/Nd}}$ in OIBs is thought to be caused by either recent metasomatism or by extremely small amounts of melting of a garnet lherzolite¹⁷. For MORBs, metasomatism is a viable, though *ad hoc*, explanation for the low Sm/Nd and Lu/Hf in the basalts; we will consider the more likely case that melting has decreased Lu/Hf and Sm/Nd in MORBs, and has led to the observed $\delta_{\text{Lu/Hf}}$ and $\delta_{\text{Sm/Nd}}$.

We used the melting equation for equilibrium non-modal partial melting to calculate the trace-element ratios in the melt. We chose the proportions in which the phases enter the melt so that the melt was basaltic in composition. Our first conclusion is that simple one-stage melting models with anhydrous spinel lherzolite as a source material cannot account for the observed $\delta_{\text{Lu/Hf}}$ in MORBs, unless extremely small amounts of melt are extracted (<0.1%). The combined $\delta_{\text{Sm/Nd}}$ and $\delta_{\text{Lu/Hf}}$ systematics cannot be satisfied with any degree of melting in the spinel stability field, using either batch equilibrium melting or any of the more complicated melting models^{1,18–22}. The Sm–Nd and Lu–Hf isotope systematics cannot be simultaneously satisfied even by using an extreme set of partition coefficients for clinopyroxene (increasing D_{Lu} so that $D_{\text{Lu}}/D_{\text{Hf}} > 3$, well above the experimental values).

A phase with $D_{\text{Lu}}/D_{\text{Hf}}$ larger than clinopyroxene but with 'comparable' D s for Sm and Nd is required. Garnet is such a phase; the relevant partition-coefficient ratios are $D_{\text{Lu}}/D_{\text{Hf}} = 23.9$ and $D_{\text{Sm}}/D_{\text{Nd}} = 4.3$ (for clinopyroxene $D_{\text{Lu}}/D_{\text{Hf}} = 1.55$ and $D_{\text{Sm}}/D_{\text{Nd}} = 1.60$). The garnet–melt distribution coefficients are from ref. 23, and our unpublished data on garnet–diopside partitioning in garnet lherzolite xenoliths, equilibrated at 19–27 kbars, from Pali-Aike, Chile²⁴. Garnet has an important role in OIB genesis; it has been proposed as a residual phase in MORB genesis^{9,25,26} and as an important phase controlling the trace elements in alpine peridotites²⁷ and abyssal peridotites²⁶. Amphibole is another phase that can change the K_{ds} appropriately. MORBs are undersaturated in H₂O, however, and thus amphibole cannot be a residual phase during the majority of the melting in MORB genesis.

† Now at: Lamont–Doherty Geological Observatory, Palisades, New York 10964, USA

‡ Now at: Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

In our melting model, which satisfies the Sm/Nd and Lu/Hf systematics, melting starts in the garnet stability field. In attempting to make the model as realistic as possible, we took the following observations into account. Recent calculations on melt extraction indicate that small amounts of melt (0.1–1%) are able to separate from the solid and ascend towards the surface¹. We modelled this sequentially such that when a critical degree of melting (0.1–1%) is reached, this melt escapes and is accumulated at the top. The next batch is generated at shallower depth, with slightly different mineralogy (and mineral chemistry) and from a source already depleted by the previous melt extraction. Although this process is perhaps continuous, we treat it as a sequential batch process for numerical simplicity. Extracting melt only above a certain threshold, as in the critical melting model²¹, would not substantially change the results.

The importance of the shape of the melting region has been stressed¹⁸. As the mantle material rises, the material at the ridge axis will end up higher than its neighbouring material that is somewhat off-axis²⁸. Furthermore, the isotherm at the bottom of the plate is undisturbed and a ridge is not dynamically compensated^{1,29}. Consequently, material directly under the ridge and material slightly off-axis will start melting at the same depth. The on-axis peridotite will rise to shallower depths and thus melt to a larger degree than the off-axis peridotite, which stops rising and 'turns the corner' at greater depth. In our model the melt generated off-axis is funnelled and collected at the centre. We assume that the decrease in degree of melting away from the ridge is linear^{28,30} (see inset in Fig. 2 for further details).

Figure 2 illustrates the effect of garnet on the Sm/Nd and Lu/Hf ratios using the above equilibrium sequential batch melting model. Garnet is residual in the deeper part of the melting column. There are two useful ways (both shown in Fig. 2) to express the amount of melting in the garnet stability field: (1) the fraction of melt, of the total, that was generated in the garnet field; (2) the fraction of the on-axis melting column that lies in the garnet field. The degree of melting as seen at the surface is the average degree of melting and, because the shape of our melting region is linear, the average degree of melting is half the amount of on-axis melting. The middle of the MORB field in Fig. 2 can be reached with 10–15% melting on-axis, with one-fifth to one-third of the on-axis melting occurring in the garnet stability field. The low $\delta_{\text{Sm}/\text{Nd}}$, high $\delta_{\text{Lu}/\text{Hf}}$ MORBs can be generated by 15% of melting on-axis with one-third of the on-axis melt generated in the garnet field. Simpler melting models with residual garnet are also able to generate the observed variations, but smaller degrees of melting are required. Whatever melting model is chosen, some garnet is always

required. Figure 2 also illustrates, as is commonly known, that many OIBs can only be generated by small degrees of melting, mostly in the garnet stability field.

Our model in Fig. 2 uses individual melt batches of 0.5%, and a garnet/diopside ratio of unity. If 0.1% of melt is allowed to escape, then slightly less on-axis melting is needed, 3–14%, to obtain the full range of $\delta_{\text{Lu}/\text{Hf}}$ and $\delta_{\text{Sm}/\text{Nd}}$ in MORBs. A maximum amount of on-axis melt (between 7–20%) can be obtained by extracting individual batches of 1%. The lower degrees of total melting needed with small (0.1%) individual batches is a consequence of the efficiency of this process, which effectively removes all the trace element from the residue.

The dynamic-melting model proposed by Langmuir *et al.*²⁰ (or the model proposed by Johnson *et al.*²⁶) is also capable of explaining the $\delta_{\text{Sm}/\text{Nd}}$ and $\delta_{\text{Lu}/\text{Hf}}$, however smaller degrees of melting are necessary. The models proposed by McKenzie and Bickle¹, or Prinzhofer and Allegre¹⁹ cannot generate the observed $\delta_{\text{Sm}/\text{Nd}}$ and $\delta_{\text{Lu}/\text{Hf}}$. The correlation between $\delta_{\text{Lu}/\text{Hf}}$ and other indicators of degree of melting, such as $\text{Na}_{8.0}$ (ref. 10), confirms the idea that $\delta_{\text{Lu}/\text{Hf}}$ can be used as an indicator for the degree of melting. Furthermore, our estimates of the degree of melting, 5–17%, agrees with the estimates of Klein and Langmuir¹⁰.

The onset of melting in the garnet stability field places the geotherm under the ridge on the solidus at $1,450 \pm 50^\circ\text{C}$ at 80–90 km (refs 17, 31). This estimate is within the uncertainty of the Parsons and Sclater³² estimate of $1,350 \pm 275^\circ\text{C}$ at 125 km for the base of the lithosphere, but their estimate seems to be somewhat low. It is difficult to calculate whether the total amount of melt produced is realistic because the behaviour of the solidus as a function of melt depletion of the peridotite is not known. Experiments on peridotites³³ indicate an increase in the solidus by $\approx 50^\circ\text{C}$ for 20% batch melting of a fertile peridotite; this is a minimum estimate for a sequentially melted source. Assuming the increase in the solidus is 75°C and using heat-of-fusion data from ref. 34, then the magma reaches $1,300$ – $1,320^\circ\text{C}$ at 20-km depth with 15% of adiabatic batch melting. Assuming the onset of melting at 90-km depth, then the relatively large degree of melting in the garnet stability field can be explained by the effect of increased amounts of melt generated at a cusp of the solidus³⁵ related to the garnet–spinel transition. Although more knowledge of the solidus of depleted peridotite is necessary to test our model fully, our conclusions are consistent with the limited information derived from experimental data.

The $\delta_{\text{Lu}/\text{Hf}}$ and $\delta_{\text{Sm}/\text{Nd}}$ calculated above assumed a depletion age of 2 Gyr for the MORB mantle. However, an average residence time of 600 Myr for some trace elements in the MORB

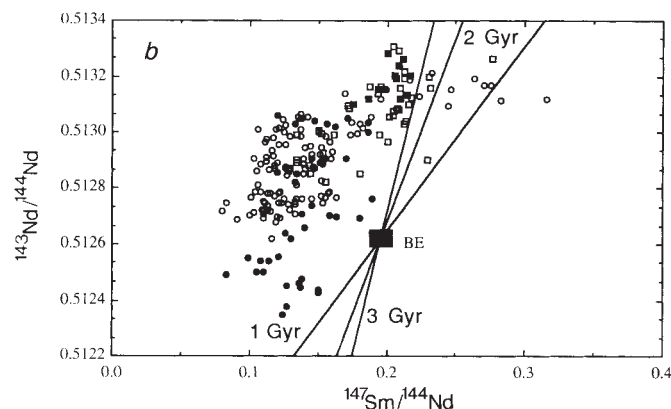
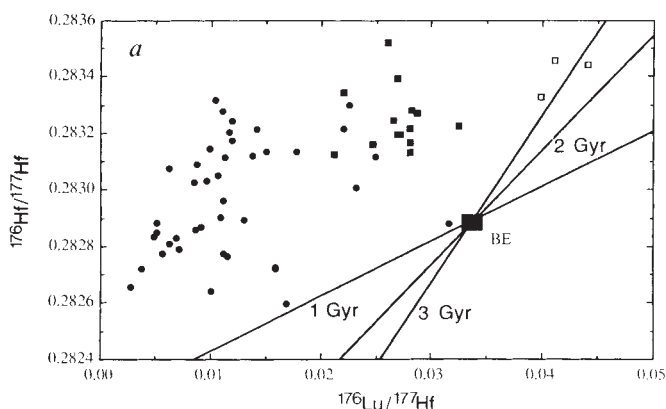


FIG. 1 a, $^{176}\text{Hf}/^{177}\text{Hf}$ plotted against $^{176}\text{Lu}/^{177}\text{Hf}$ and b, $^{143}\text{Nd}/^{144}\text{Nd}$ plotted against $^{147}\text{Sm}/^{144}\text{Nd}$ for oceanic volcanics. Squares are MORBs and circles are OIBs. Solid symbols are for those samples where both the Lu–Hf and Sm–Nd systematics are known. The 1-, 2- and 3-Gyr lines are isochrons through BE (bulk Earth). The geochron is a near-vertical line through BE (not

shown). Samples falling to the left of the isochrons must have undergone a recent decrease in parent/daughter ratio. Data from refs 5, 6, 37 and references therein, W. M. White, unpublished data, and MIT, unpublished data. Bulk Earth Lu/Hf from ref. 38 and C1 (or BE) $^{176}\text{Hf}/^{177}\text{Hf}$ from ref. 6.

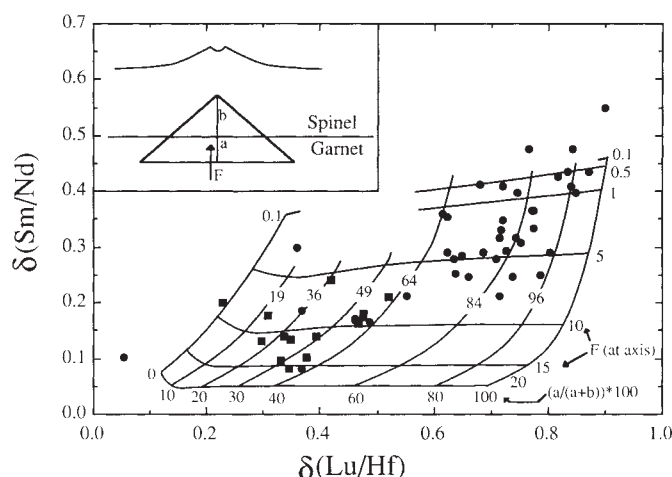


FIG. 2 $\delta_{Lu/Hf}$ plotted against $\delta_{Sm/Nd}$ for MORBs and OIBs. Symbols for basalts and data sources as in Fig. 1. $\delta_{Lu/Hf}$ and $\delta_{Sm/Nd}$ are the fractional differences between parent/daughter measured in the basalt and the calculated parent/daughter ratio if this sample would fall on the 2-Gyr isochron. The curves of the non-modal sequential batch equilibrium melting model are for different degrees of melting and for varying amounts of melting in the garnet stability field. Inset in the diagram is a representation of the melting model. The vertical axis is the degree of melting (F) in a column of mantle, that is, the total amount of melt generated by a unit volume of peridotite. $a/(a+b)$ is the percent of melt in a melt column generated in the garnet field; this number appears as the bottom row of numbers below the melting curves. Numbers on right of the diagram decreasing in the y -direction are the different degrees of melting at the ridge axis, that is, the maximum degree of melting ($a+b$). Initial garnet/diopside ratio in the peridotite is unity. Numbers on the curves are the percentage of the total amount of melt collected as the 'top' of melt, that was generated in the garnet stability field. In the inset this is the area within the triangle and below the spinel-garnet transition as percentage of the total area of the triangle. The breakdown of 1% garnet produces 0.5% of diopside. Our preferred set of D s for olivine, orthopyroxene and clinopyroxene are taken from those listed or referred to in refs 21, 39, 40. The distribution coefficients for Sm, Nd, Lu and Hf are 0.37, 0.26, 0.45 and 0.37 respectively for diopside, and 0.13, 0.03, 5.5 and 0.23 respectively for garnet. Although the absolute D s of the elements might not be known accurately we think that for the rare-earth elements and Hf the d ratios are better constrained. Therefore our modelling only involves modelling of trace-element ratios without emphasis on the trace-element abundances. In the garnet stability field the mode of a phase in the melt is 0.45, 0.45 and 0.1 for garnet, diopside and olivine + orthopyroxene (+spinel) respectively. In the spinel stability field the mode in the melt for clinopyroxene and olivine + orthopyroxene is 0.7 and 0.3 respectively. Starting composition is 80% olivine + orthopyroxene (+spinel), 10% garnet and 10% clinopyroxene.

reservoir has been suggested³⁶, based on U-Th-Pb trace-element and isotope systematics. Decreasing the age of the MORB reservoir to 600 Myr will increase the $\delta_{Lu/Hf}$ and $\delta_{Sm/Nd}$; this will require an increase in the proportion of melting in the garnet field (larger $\delta_{Lu/Hf}$ requires smaller F and larger $a/a+b$, see Fig. 2). The U-Th-Pb arguments are partly based on the partitioning behaviour of these trace elements, however, which is unknown in the presence of garnet. Partition studies for U-Th and Pb between garnet and melt are necessary in order to evaluate the thorium-isotope constraints on MORB mantle processes.

The Hf-Nd isotope and trace-element characteristics of MORBs provide strong constraints on the depth of origin of MORB and unquestionably require garnet as a residual phase during the early stages of melting. Future ridge models will have to start melting at depths greater than 90 km. Furthermore to satisfy the Sm-Nd and Lu-Hf constraints, either sequential melting or very low degrees (0.5–2%) of batch melting is required. We prefer a sequential melting model because this model allows higher degrees of total melt and because the rare-earth-element variations in abyssal peridotites²⁶ support such a model. □

Received 14 July; accepted 16 October 1989.

- McKenzie, D. & Bickle, M. J. *J. Petrol.* **29**, 625–679 (1988).
- Falloon, T. J. & Green, D. H. *Miner. Petrol.* **37**, 181–219 (1987).
- Elthon, D. *Crit. Rev. Mar. Sci.* (in the press).
- Salter, V. J. M. & Hart, S. R. *Trans. Am. geophys. Un.* **70**, 510 (1989).
- Patchett, P. J. *Geochim. cosmochim. Acta* **47**, 81–91 (1983).
- Stille, P., Unruh, D. M. & Tatsumoto, M. *Nature* **304**, 25–29 (1983).
- White, W. M. & Patchett, P. J. *Earth planet. Sci. Lett.* **67**, 167–185 (1984).
- Dick, H. J. B., Fisher, R. L. & Bryan, W. B. *Earth planet. Sci. Lett.* **69**, 88–106 (1984).
- Bender, J. F., Langmuir, C. H. & Hanson, G. N. *J. Petrol.* **25**, 213–254 (1984).
- Klein, E. M. & Langmuir, C. H. *J. geophys. Res.* **92**, 8089–8115 (1987).
- Stolper, E. *Contrib. Miner. Petrol.* **74**, 13–27 (1980).
- Wood, D. A. *Geology* **7**, 499–503 (1979).
- Thompson, R. N., Morrison, M. A., Dickin, A. P. & Hendry, G. L. in *Continental Basalts and Mantle Xenoliths* (eds Hawkesworth, C. J. & Norry, M. J.) (Shiva, Cheshire, 1983).
- Zindler, A. & Hart, S. R. *Rev. Earth planet. Sci.* **14**, 493–571 (1986).
- Allegre, C. J., Hart, S. R. & Minster, J. F. *Earth planet. Sci. Lett.* **66**, 191–213 (1983).
- DePaolo, D. J. & Wasserburg, G. J. *Geophys. Res. Lett.* **3**, 743–746 (1976).
- Basaltic Volcanism Study Project. *Basaltic Volcanism on the Terrestrial Planets* (Pergamon, New York, 1981).
- O'Hara, M. J. *Nature* **314**, 58–62 (1985).
- Prinzhofer, A. & Allegre, C. J. *Earth planet. Sci. Lett.* **74**, 251–265 (1985).
- Langmuir, C. H., Bender, J. F., Bence, A. E., Hanson, G. N. & Taylor, S. R. *Earth planet. Sci. Lett.* **36**, 133–156 (1977).
- Maaloe, S. *Geochim. cosmochim. Acta* **46**, 43–57 (1982).
- Navon, O. & Stolper, E. *J. Geol.* **95**, 285–307 (1987).
- Fujimaki, H., Tatsumoto, M. & Aoki, K. *J. geophys. Res.* **89**, 662–672 (1984).
- Stern, C. R., Saul, S., Skewes, M. A. & Futa, K. *Geol. Soc. Aust. Spec. Publ.* **14**, 735–744 (1989).
- O'Hara, M. J. *Earth. Sci. Rev.* **4**, 69–133 (1968).
- Johnson, K. T. M., Dick, H. J. B. & Shimizu, N. *J. geophys. Res.* (submitted).
- Loubet, M., Shimizu, N. & Allegre, C. J. *Contrib. Miner. Petrol.* **53**, 1–12 (1975).
- Phipps Morgan, J. *Geophys. Res. Lett.* **14**, 1238–1241 (1987).
- Richter, F. M. & McKenzie, D. *J. geophys. Res.* **86**, 1738–1744 (1981).
- Scott, D. R. & Stevenson, D. J. *J. geophys. Res.* **94**, 2973–2988 (1989).
- Scarf, C. M. & Brearley, M. in *Mantle Xenoliths* (ed. Nixon, P. H.) 599–608 (Wiley, New York, 1987).
- Parsons, B. & Sclater, J. G. *J. geophys. Res.* **82**, 803–827 (1977).
- Green, D. H., Hiberson, W. O. & Jacques, A. L. in *The Earth: Its Origin Structure and Evolution* (ed. McElhinney, M. W.) 265–299 (Academic, London, 1979).
- Ghiorsio, M. S. *Contrib. Miner. Petrol.* **90**, 107–120 (1985).
- Presnall, D. C. & Hoover, J. D. in *Magmatic Processes: Physicochemical Principles* (ed. Mysen, B. O.) 75–89 (Geochem. Soc., 1987).
- Galer, S. J. G. & O'Nions, R. K. *Nature* **316**, 778–782 (1985).
- Hart, S. R. & Zindler, A. in *Mantle Convection* (ed. Peltier, D.) 261–387 (1988).
- Beer, H., Walter, G., Macklin, R. L. & Patchett, P. J. *Phys. Rev. C30*, 464–478 (1984).
- Irving, A. J. & Frey, F. A. *Geochim. cosmochim. Acta* **48**, 1201–1221 (1984).
- Frey, F. A., Green, D. H. & Roy, S. D. *J. Petrol.* **19**, 463–513 (1978).

ACKNOWLEDGMENTS. We thank W. White for the use of unpublished hafnium isotope data, J. Blustzjan for some analyses, T. Grove, J. Phipps Morgan and T. Juster for suggestions and discussions and N. Shimizu for discussions and for reading the manuscript.

A new aluminocalcic high-pressure phase as a possible host of calcium and aluminium in the lower mantle

M. Madon, J. Castex & J. Peyronneau

Département des Géomatériaux, Institut de Physique du Globe de Paris, 4 Place Jussieu, 75252 Paris cedex 05, France

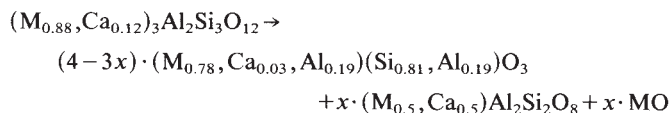
IT IS generally agreed that most silicates in the Earth's lower mantle have the perovskite structure, with the unit cell being orthorhombic for small cations (such as magnesium and aluminium) and cubic for larger cations (such as calcium). To study the possibility of alternative host minerals for the large cations in the deep mantle, we have transformed natural ($M_{0.88}$, $Ca_{0.12}$) $Al_2Si_3O_{12}$ garnets (with $M = Mg, Fe$ or Mn) at high pressure (50 GPa) and high temperature in a laser-heated diamond-anvil cell. The quenched products were studied by analytical transmission electron microscopy; this revealed the presence of a new phase with a composition close to $(Ca_{0.5}, M_{0.5})Al_2Si_2O_8$. High-pressure experiments performed with a glass of similar composition $((Ca_{0.5}, Mg_{0.5})Al_2Si_2O_8)$ confirm the existence of this new phase at a pressure of 50 GPa. The cell dimensions and stoichiometry of this new phase show strong similarities to those of the hollandite structure. Depending on the chemical composition of the lower mantle, this aluminocalcic phase could be the second or the third most abundant constituent of the lower mantle, in which case the abundance of $CaSiO_3$ perovskite would be negligible.

Any mantle model¹⁻⁴ includes significant amounts of both CaO and Al₂O₃, in addition to SiO₂, MgO and FeO. From high-pressure experiments performed on simple chemical systems, it has been suggested that, in the deep mantle, calcium is found in a CaSiO₃ perovskite^{5,6}, with aluminium in an Al-bearing perovskite^{7,8}. Several experiments with more complex chemical systems have shown, however, that other aluminocalcic high-pressure phases could be formed. For instance, a CaO-rich phase and an Al₂O₃-rich phase have been observed in experiments performed at 27 GPa and 1,600 °C, the composition of the starting material being close to the olivine-subtracted peridotitic mantle composition^{9,10}. Owing to the small grain size, the exact composition and the structure of these new phases are unknown.

To identify such aluminocalcic high-pressure phases, we have transformed a natural garnet in a laser-heated diamond-anvil cell. The starting material was a xenocryst, from Moses Rock (Arizona), with a composition close to (M_{0.88}, Ca_{0.12})₃Al₂Si₃O₁₂ where M corresponds to divalent cations Mg, Fe and Mn (see Table 1). The starting material was crushed in an agate mortar, and the powder was loaded in the hole of a nickel gasket. Samples were then brought to ~50(±5) GPa, as determined with a strain gauge¹¹, and heated by scanning the beam of a Nd YAG laser across the sample. After heating, the quenched products were ion-beam thinned and observed with a Jeol 2000 EX electron microscope equipped with a Tracor TN 5400 FX analytical attachment. The transformed materials were identified by electron diffraction and/or microanalysis. All experimental procedures, and especially the precision of chemical analysis, are described elsewhere^{11,12}.

In all samples, observations by transmission electron microscope (TEM) show the coexistence of two phases. The first is perfectly stable under the electron beam, and it is found in contact with another phase that amorphizes quickly under the

electron beam (Fig. 1a). The electron diffraction patterns taken on this latter phase before amorphization are compatible with the structure of orthorhombic perovskite¹³, but those obtained on the stable phase are new: they do not correspond to a classical high-pressure structure such as perovskite. The composition of the two phases, given by *in situ* X-ray microanalysis, are close to (M_{0.78}, Ca_{0.03}, Al_{0.19})(Si_{0.81}, Al_{0.19})O₃ for the perovskite phase and (M_{0.5}, Ca_{0.5})Al₂Si₂O₈ for the other phase. The composition of the starting material and those of the final products (Table 1), indicate that the decomposition of garnet must also lead to the formation of a magnesiowüstite, according to the reaction:



where *x* is a factor depending of the initial content of calcium in the starting material. In our experiments (Ca/Ca + M = 0.12), *x* = 0.557.

No magnesiowüstite was observed in the samples, but this could be due to the small amount of this phase formed during the decomposition of garnet into post-garnet phases. Indeed, from the molar volume of each component, ~6% of magnesiowüstite should be formed in the previous reaction, for 56% of perovskite and 38% of the new phase. The relative proportions of perovskite and new phase observed in the transformation products of garnet are in good agreement with these values.

No CaSiO₃ perovskite was observed in the transformation products of garnet. We conclude that this phase cannot be formed at high pressure (at least at 50 GPa), when the three types of cations—calcium, aluminium and magnesium (or iron)—exist simultaneously in the starting material.

To see if a change in the chemical composition of the starting material leads to the same aluminocalcic phase, and to determine

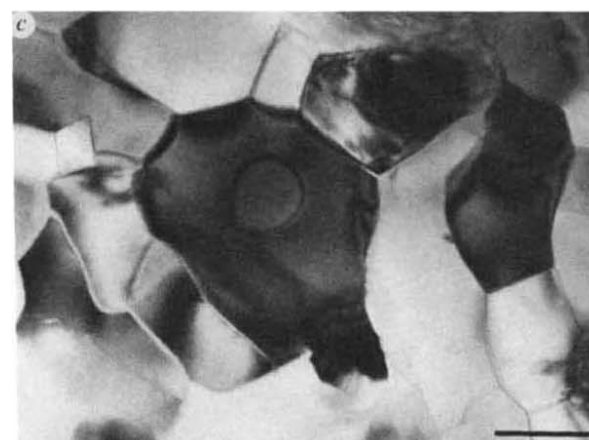
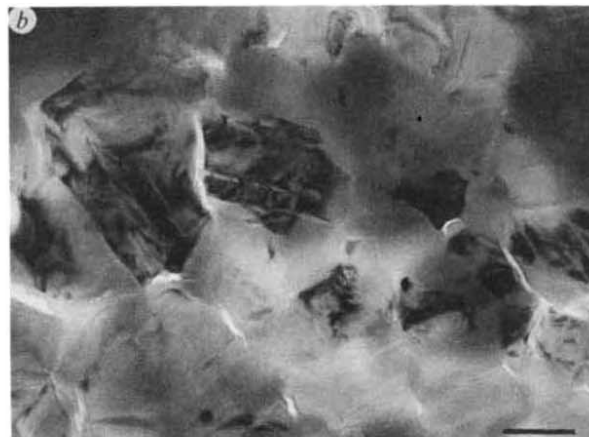
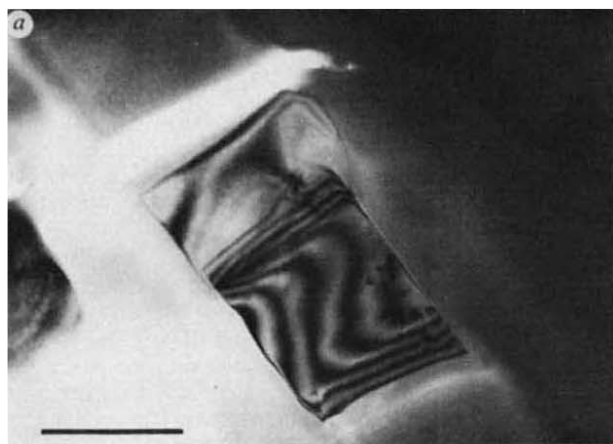


FIG. 1 Transmission electron micrographs. a, The two high-pressure phases identified in samples of natural garnet brought to 50 GPa in a diamond-anvil cell, and heated by a laser beam. The perovskite (white) has become amorphous in the electron beam. It is found in contact with a crystalline phase of composition close to (M_{0.5}, Ca_{0.5})Al₂Si₂O₈, with M=Mg+Fe. Scale bar, 0.5 μm. b, The decomposition of anorthite CaAl₂Si₂O₈ at high pressure (50 GPa) into a glass of composition CaSiO₃ and a crystalline phase of composition Al₂SiO₅. The glass is assumed to be the result of the amorphization of the CaSiO₃ perovskite, and the crystalline phase is unidentified. Scale bar, 0.5 μm. c, The high-pressure phase of a glass of composition (Mg_{0.5}, Ca_{0.5})Al₂Si₂O₈. Scale bar, 0.5 μm.

the cell parameters, we performed additional runs with two glasses as starting materials. The first glass had the composition of anorthite [$\text{CaAl}_2\text{Si}_2\text{O}_8$] and the second glass, the same composition with half atoms of magnesium in substitution of atoms of calcium [$(\text{Ca}_{0.5}, \text{Mg}_{0.5})\text{Al}_2\text{Si}_2\text{O}_8$]; the latter composition is very close to that described in the previous section. The finely powdered glasses, mixed with a trace amount of platinum to absorb the laser radiation, were compressed and heated in the diamond-anvil cell at the same (P , T) conditions than those previously described. The TEM observations of the transformation products of two glasses lead to the following comments: (1) At high pressure, $\text{CaAl}_2\text{Si}_2\text{O}_8$ decomposes in two phases, a

TABLE 1 Chemical composition (in atom %) of the starting material (garnet) and of the three post-garnet components

Element	Garnet*	Perovskite†	(Ca,Mg)-phase‡	Magnesiowüstite‡
Si	15.00	16.20	15.18	00.00
Ti	00.03	00.02	00.02	00.00
Ca	01.80	00.57	04.05	00.00
Mg	10.05	13.35	03.24	20.64
Fe	03.04	02.12	00.82	28.24
Mn	00.11	00.08	00.01	01.12
Al	09.66	07.35	15.05	00.00
Cr	00.26	00.25	00.25	00.00
O§	60.00	59.96	61.47	50.00

* Analyses from a Camebax (Camparis, Université Paris 6) electron microprobe.

† Analyses from the Tracor TN5400 energy-dispersive X-ray analyser.

‡ Estimated composition of magnesiowüstite (see text).

§ Oxygen given by stoichiometry.

TABLE 2 X-ray powder diffraction data for $(\text{Ca}_{0.5}, \text{Mg}_{0.5})\text{Al}_2\text{Si}_2\text{O}_8$ phase

$h\ k\ l$	monoclinic	d_{obs}	d_{calc}	I/I_0
1 2 0		3.738	3.736	8
1 -2 -1		3.467	3.469	10
3 0 0		3.130	3.128	17
1 0 -3		2.937	2.939	26
0 3 0		2.702*	2.716	29
0 3 -1		2.607	2.606	63
2 1 -3		2.468	2.467	47
0 3 -2		2.341	2.343	7
1 0 4		2.243	2.243	94
3 0 3		2.189	2.188	33
0 3 3		2.040	2.039	100
1 2 -4		1.972	1.971	25
3 2 -3		1.941	1.940	43
4 2 -2		1.866	1.866	9
2 0 -5		1.728	1.727	12
1 2 5		1.658	1.657	5
5 0 3		1.599	1.598	23
6 0 0		1.563	1.564	57
4 4 0		1.537	1.538	40
0 1 6		1.516	1.516	6
2 0 -6		1.469	1.469	<5
2 1 6		1.439	1.439	<5
6 0 -3		1.400	1.400	<5
4 -4 3		1.373	1.374	21
1 6 0		1.345	1.344	34
3 5 3		1.307	1.307	<5
1 6 3		1.232	1.232	<5
2 6 -3		1.202	1.202	<5
4 6 0		1.175	1.175	22
1 0 8		1.147	1.147	<5
7 3 -3		1.123	1.123	6

* Reflection not used for lattice-constant refinement.

Calculated from $a=9.384(2)$, $b=8.148(2)$, $c=9.258(2)$ Å and $\beta=90.474(22)^\circ$. $V=707.89$ Å³ (71.06 cm³ mol⁻¹), $\rho=3.804$ g cm⁻³.

glass of composition CaSiO_3 , in contact with another crystalline phase of composition Al_2SiO_5 (Fig. 1b). The amorphous CaSiO_3 phase is assumed to be the result of the amorphization of CaSiO_3 perovskite at room pressure, as generally observed¹⁴, and the crystalline Al_2SiO_5 phase is unidentified, but the electron diffraction patterns disagree with the kyanite structure; (2) The experiments done with the second glass confirm the existence of a crystalline phase of $(\text{Ca}_{0.5}, \text{Mg}_{0.5})\text{Al}_2\text{Si}_2\text{O}_8$ composition, at 50 GPa (Fig. 1c). These experiments confirm the stability of the CaSiO_3 perovskite at high pressure when two types of cations (Ca and Al) exist in the starting material, but also show that another high-pressure phase is preferred when the three cations (Ca, Al and Mg or Fe) are simultaneously present in the sample.

The cell parameters of the new phase were identified from powder X-ray diffraction patterns obtained with a 180-mm Debye-Scherrer camera. X-ray diffraction data for a glass sample of $(\text{Ca}_{0.5}, \text{Mg}_{0.5})\text{Al}_2\text{Si}_2\text{O}_8$, quenched from ~50 GPa, are listed in Table 2. The unit-cell dimensions were determined to be $a=9.384$, $b=8.148$ and $c=9.258$ Å, with $\beta=90.474^\circ$. The molar volume and the density of the new phase calculated for $Z=6$, are 71.06 ± 0.04 cm³ mol⁻¹ and 3.804 ± 0.002 g cm⁻³ respectively, at room temperature and pressure. This phase is very similar to the hollandite structure. First, the molar volume and the density are similar to those of $\text{NaAlSi}_3\text{O}_8$ hollandite¹⁵ ($V=71.0$ cm³ mol⁻¹, $\rho=3.69$ g cm⁻³) and KAlSi_3O_8 hollandite¹⁶ ($V=72.6$ cm³ mol⁻¹, $\rho=3.84$ g cm⁻³). Second, the composition of the (Ca,Mg) phase conforms to the general formula $[\text{A}_x][\text{B}_y, \text{C}_{4-y}]\text{O}_8$ of the hollandite structure family where the A cations may be monovalent or divalent, whereas the smaller cations making up the octahedral $[\text{B,C}]_4\text{O}_8$ framework include tri- and tetravalent species¹⁷. Third, the lattice parameters are very close to those of the distorted hollandite structure. The ideal hollandite unit cell is tetragonal, with a value of the 'c' axis ~2.8 Å. This value represents the edge-to-edge dimension of a $[\text{B,C}]_6$ octahedra. Yet, depending on the chemistry of the sample, the c axis can be a multiple of the ideal value. This is the case in many barium titanate hollandites where a deficiency in stoichiometry leads to a multiplicity of the c axis¹⁸. On the other hand, many hollandites show a small distortion that lowers the symmetry from the tetragonal to the monoclinic structure, with a value of $\beta \sim 90^\circ$ (refs 19, 20). In our experiments the aluminocalcic phase seems to adopt such a distorted hollandite structure, that is a monoclinic structure ($\beta=90.474^\circ$) with a multiplicity of the b axis (equivalent to the c axis of the ideal structure) close to three. But two points disagree with this interpretation. First, the multiplicity of the b axis is not easily explained because the (Ca,Mg) phase seems to be stoichiometric. Second, the intensities of the diffraction lines (Table 2) are not in good agreement with those of the ideal hollandite structure. Further work is needed to resolve the issue.

The density of the $(\text{Ca}_{0.5}, \text{Mg}_{0.5})\text{Al}_2\text{Si}_2\text{O}_8$ phase seems to be low at room pressure ($\rho=3.804$ g cm⁻³) with respect to the densities of other high-pressure phases such as MgSiO_3

TABLE 3 Estimates of mineralogical constitution of the lower mantle (in volumetric ratio) for different models of chemical composition

Composition model	Pyrolite*	Lherzolite†	Chondritic†	Piclogite†
Perovskite				
($\text{M}_{0.97}, \text{Ca}_{0.03}$) SiO_3	65.0	70.5	76.6	55.2
Magnesiowüstite				
MO	20.0	23.0	11.6	9.2
(Ca,Mg)-phase				
($\text{M}_{0.5}, \text{Ca}_{0.5}$) $\text{Al}_2\text{Si}_2\text{O}_8$	12.7	6.0	10.5	25.6
Perovskite				
CaSiO_3	2.3	0.5	1.3	10.0

Abundance of elements in each composition model is described elsewhere^{4,27}. M=Mg+Fe.

* See ref. 27.

† See ref. 4.

perovskite^{21,22} ($\rho = 4.11 \text{ g cm}^{-3}$), or CaSiO_3 perovskite^{23,24} ($\rho = 4.24 \text{ g cm}^{-3}$). As $(\text{Ca}_{0.5}, \text{Mg}_{0.5})\text{Al}_2\text{Si}_2\text{O}_8$ forms at 50 GPa along with the magnesian perovskite, we can expect that its density is similar to that of perovskite MgSiO_3 (or CaSiO_3) at the same pressure. At 50 GPa and room temperature, the perovskite densities are $\sim 4.9 \text{ g cm}^{-3}$. To obtain the same density at the same pressure, the $(\text{Ca}_{0.5}, \text{Mg}_{0.5})\text{Al}_2\text{Si}_2\text{O}_8$ phase should have a bulk modulus of $\sim 120 \text{ GPa}$ with the commonly made assumption ($K'_0 = 4$) for the first derivative of K_0 at 1 bar. We note that this value is low in comparison with those of perovskites ($K_0 = 247 \text{ GPa}$ in MgSiO_3 perovskite²² and 268 GPa in CaSiO_3 perovskite²⁴), but of the same order of magnitude as other compounds such as wüstite MgO (ref. 25) ($K_0 = 154 \text{ GPa}$) or CaO (ref. 26) ($K_0 = 111 \text{ GPa}$). Another explanation for the low value of the density of the aluminocalcic phase at room pressure could be a phase change of this compound at higher pressure—indeed, a reversible transition (similar to the B1–B2 transition occurring in the CaO compound at ~ 60 or 70 GPa (ref. 26) could lead the $(\text{Ca}_{0.5}, \text{Mg}_{0.5})\text{Al}_2\text{Si}_2\text{O}_8$ phase to a density similar to those of perovskites at high pressure and to a low density at room pressure.

We have started static measurements of the compression of the $(\text{Ca}_{0.5}, \text{Mg}_{0.5})\text{Al}_2\text{Si}_2\text{O}_8$ phase in a diamond-anvil cell to determine its high-pressure density and look for a possible phase transition.

These observations show that an aluminocalcic phase with a composition $[(\text{Ca}_{0.5}, \text{Mg}_{0.5})\text{Al}_2\text{Si}_2\text{O}_8]$ close to but different from that of anorthite, can accommodate the calcium at high pressure. We suggest that this phase could be the crystalline host for not only the calcium in the lower mantle, but also for the other large cations such as sodium, potassium or the radioactive elements. This phase could be an important constituent of the lower mantle. Depending on the assumed mantle composition, this phase is inferred to be the second or the third most abundant constituent of the lower mantle, with the perovskite and the magnesiowüstite (Table 3). Indeed, assuming that all the aluminium reacts with the calcium to form an aluminocalcic phase of the above composition, and that the solubility of the CaSiO_3 molecule in MgSiO_3 orthorhombic perovskite is $\sim 3\%$, the proportion of the (Ca, Mg) -phase, in volumetric ratio, would vary from 6% to 25% of the lower mantle. As a result, the abundance of CaSiO_3 perovskite in the lower mantle would be negligible. $\square$

Discovery of the earliest-known tetrapod stapes

J. A. Clack

University Museum of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK

THE evolution of the middle ear is central to the discussion of how the first tetrapods adapted to life on land as well as their phylogeny^{1–3}. Here I report the discovery of the stapes of *Acanthostega gunnari*, from the Upper Devonian of east Greenland. This is the earliest tetrapod stapes so far described, and it throws new light on both these aspects of early tetrapod biology. It has been assumed that the common inheritance of all early tetrapods was a light, rod-like stapes associated with a temporal notch in the otic region that was thought to have supported a tympanum, or eardrum. The stapes would have conducted vibrations from the tympanum to the otic capsule. By contrast, the stapes of *Acanthostega* was stout with a broad distal ramus associated with the temporal notch. I suggest that the temporal notch of *Acanthostega* and other early tetrapods supported a spiracular opening rather than a tympanum, and that the stapes controlled palatal and spiracular movements in ventilation.

Of six *Acanthostega* stapes collected by a Cambridge–Copenhagen expedition to east Greenland (Figs. 1, 2; ref. 4), four are preserved approximately in life-position. This indicates that the stapes was firmly held in place in life, despite the incomplete ossification of the otic region of the braincase into which the footplate of the stapes was presumably inserted. The footplate, the medial surface of the stapodial head, is large. Only one head

Received 10 July; accepted 4 October 1989.

- Ringwood, A. E. *Composition and Petrology of the Earth's mantle*, (McGraw-Hill, New York, 1975).
- Maaloe, S. & Aoki, K. *Contr. Miner. Petrol.* **63**, 161–173 (1977).
- Anderson, D. L. in *High Pressure Research in Geophysics* (eds Akimoto, S. & Manghnani, M. H.) 301–308 (1982).
- Anderson, D. L. & Bass, D. J. *Nature* **320**, 321–328 (1986).
- Xiong, D. H., Ming, L. C. & Manghnani, M. H. *Phys. Earth planet. Inter.* **43**, 244–252 (1986).
- Tamai, H. & Yagi, T. *Phys. Earth planet. Inter.* **54**, 370–377 (1989).
- Jeanloz, R. & Thompson, A. B. *Rev. Geophys. Space Phys.* **21**, 51–74 (1983).
- Irfune, T. & Ringwood, A. E. *Earth planet. Sci. Lett.* **86**, 365–376 (1987).
- Ito, E. & Takahashi, E. in *High Pressure Research in Mineral Physics* (eds Manghnani, M. H., Syono, Y., Ito, E. & Takahashi, E.) 221–229 (1987).
- Takahashi, E. & Ito, E. in *High Pressure Research in Mineral Physics* (eds Manghnani, M. H., Syono, Y., Ito, E. & Takahashi, E.) 427–437 (1987).
- Madon, M., Guyot, F., Peyronneau, J. & Poirier, J. P. *Phys. Chem. Miner.* **16**, 320–330 (1989).
- Guyot, F., Madon, M., Peyronneau, J. & Poirier, J. P. *Earth planet. Sci. Lett.* **90**, 52–64 (1988).
- Yagi, T., Mao, H. K. & Bell, P. M. *Phys. Chem. Miner.* **3**, 97–110 (1987).
- Liu, L. G. & Ringwood, A. E. *Earth planet. Sci. Lett.* **14**, 1079–1082 (1975).
- Liu, L. G. *Earth planet. Sci. Lett.* **37**, 438–444 (1978).
- Ringwood, A. E., Reid, A. F. & Wadsley, A. D. *Acta crystallogr.* **23**, 1093–1095 (1967).
- Kesson, S. E. & White, T. J. *J. Solid St. Chem.* **63**, 122–125 (1986).
- Kesson, S. E. & White, T. J. *Proc. R. Soc. A* **408**, 295–319 (1986).
- Post, J. E., Von Dreele, R. B. & Buseck, P. R. *Acta crystallogr.* **B38**, 1056–1065 (1982).
- Mukherjee, B. *Acta crystallogr.* **13**, 164–165 (1960).
- Ito, E. & Matsui, Y. *Earth planet. Sci. Lett.* **38**, 443–450 (1978).
- Kudoh, Y., Ito, E. & Takeda, H. *Phys. Chem. Miner.* **14**, 350–354 (1987).
- Mao, H. K. *et al.* *J. geophys. Res.* (in the press).
- Tarrida, M. & Richet, P. *Geophys. Res. Lett.* (in the press).
- Richet, P., Mao, H. K. & Bell, P. M. *J. geophys. Res.* **94**, 3037–3045 (1989).
- Richet, P., Mao, H. K. & Bell, P. M. *J. geophys. Res.* **93**, 15279–15288 (1989).
- Jackson, I. *Earth planet. Sci. Lett.* **62**, 91–103 (1983).

ACKNOWLEDGEMENTS. We thank J. P. Poirier, P. Richet and A. E. Ringwood for discussions and criticism. This work was supported by CNRS (UA 734). This is Institut de Physique du Globe contribution 1095.

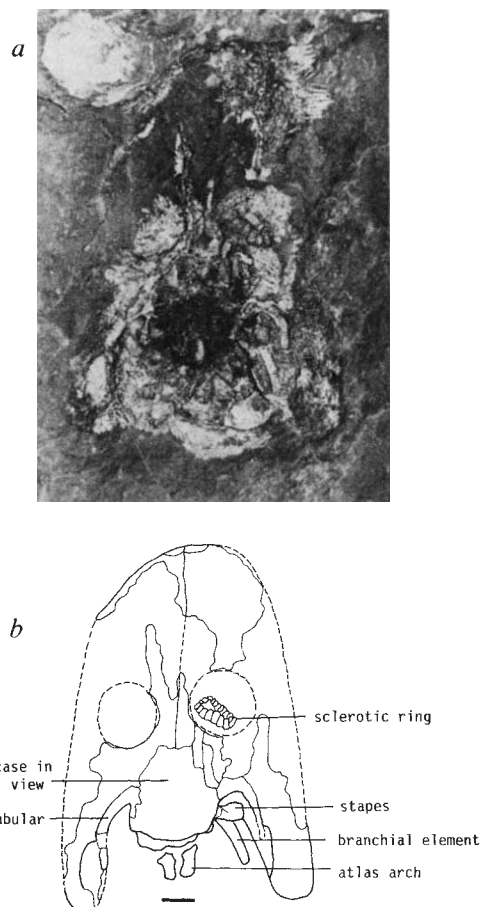


FIG. 1 *Acanthostega*, specimen 1227: a, photograph of skull in dorsal view with b, interpretive drawing. Scale bar, 10 mm.

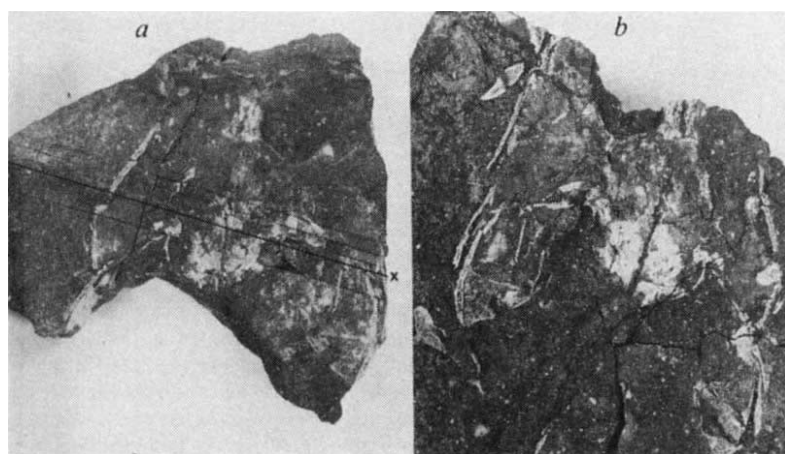


FIG. 2 *Acanthostega*, specimen 1604: a, partial skull in dorsal view showing proximal end of stapes; b, counterpart skull roof showing distal end of

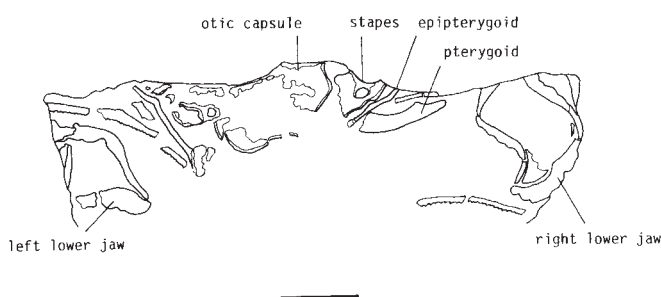


FIG. 3 *Acanthostega*, specimen 1604: drawing of section at X (see Fig. 2), showing stapes, palatal bones and lower jaws. Scale bar, 10 mm.

is apparent in vertical section, although this cannot be confirmed in the surface-prepared material, because the footplate is obscured (Figs. 3 and 4). A robust neck region separates the head from the shaft. Proximally, the head abuts the otic capsule, and the distal margin of the fan-shaped ramus, which has cartilage-finished lateral and posterior margins, extends laterally towards the well-developed temporal notch bordered by the tabular bone (Fig. 4). Temporal notches are often described as 'otic' with the presumption that each held a tympanum. Their occurrence in the earliest tetrapods has been equated with presence of a tympanum in these animals.

The massive, expanded stapes of *Acanthostega* is strikingly

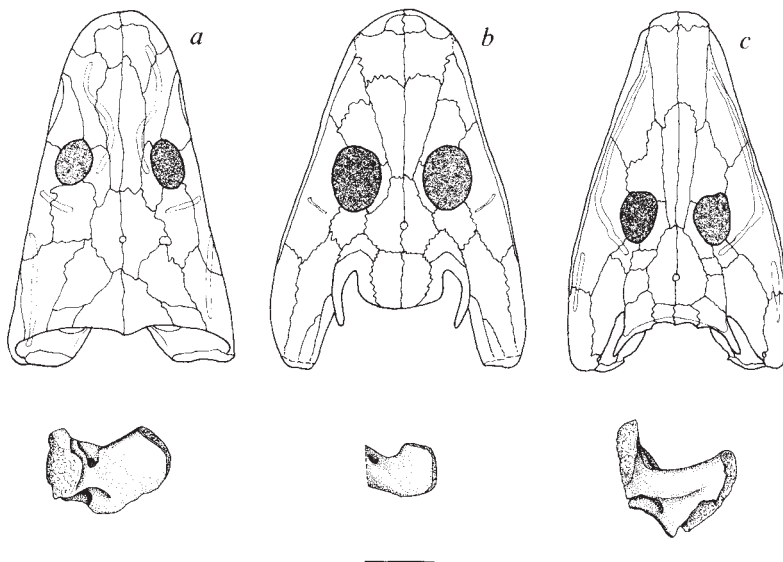
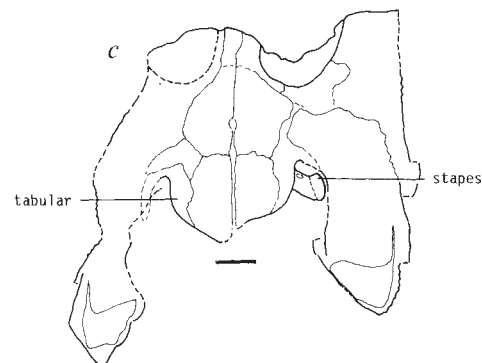


FIG. 4 Skulls and stapes of early tetrapods: a, *Greererpeton* (snout to quadrate length 131 mm), stapes in dorsal view below; b, *Acanthostega* (snout to quadrate length 115 mm), restoration of stapes from specimens 1227 and 1604 in approximately dorsal view below; c, *Pholiderpeton* (snout to quadrate length 320 mm), stapes in approximately dorsal view below. Scale bar (for stapes only), 10 mm.



stapes; c, interpretive drawing with stapes drawn from part and counterpart of specimen. Scale bar, 10 mm.

similar to those of two unrelated Carboniferous tetrapods, although it is much older^{5,6}. *Greererpeton*, a colosteid, lacks a temporal notch^{5,7}, whereas *Pholiderpeton*, an embolomere, has a narrow, wedge-shaped notch (Fig. 4). The apparent absence of a tympanic ear in these Carboniferous forms could have been a convergent loss related to a secondarily aquatic lifestyle⁶. But the stapedial morphology of *Acanthostega* indicates that it too lacked a tympanic ear, though it possessed a notch. Absence of a tympanum now appears primitive for tetrapods; its presence could only reasonably be inferred if a temporal notch were associated with a relatively light, rod-like stapes. However, even in the absence of a tympanum, the stapes is likely to have had some auditory function because of the close association between the footplate and the otic capsule⁸.

Carroll⁷ suggested that the stapes of *Greererpeton* braced the braincase against the cheek, representing the primitive tetrapod condition in which the occiput, braincase and skull roof were not fully integrated. But the braincase and skull roof were intimately connected in both *Pholiderpeton*⁹ and *Acanthostega*¹⁰, so that a bracing function for the stapes is unlikely in these animals. I suggest instead that the stapes of primitive tetrapods was actively involved in ventilation, as was its homologue, the hyomandibula of fishes.

In many fossil fishes, including the osteolepiforms usually regarded as the closest relatives of tetrapods, the hyomandibula contributed to the suspension of the lower jaw and supported the back of the palate. By analogy with living fishes, it acted as

a hinge coordinating kinetic movements of the palate and cheek used to ventilate the gills and in feeding. A similar buccal pumping mechanism is used to ventilate both lungs and gills in those fishes that have lungs¹¹. These movements are controlled by muscles attached to the palate and hyomandibula, and in the modern but primitive actinopterygian *Polypterus*, these muscles also control the operation of an open spiracle¹².

The earliest tetrapods may have retained a fish-like breathing mechanism. In *Acanthostega* and other primitive tetrapods the palate and braincase were connected by a synovial joint^{9,13} allowing a degree of intracranial mobility. The temporal notch in some primitive tetrapods may have housed an open spiracle homologous with the spiracular notch of the earliest bony fishes¹⁴. The spiracle in tetrapods could have been used for inspiration when the rest of the animal's head was submerged. The spiracular cleft and stapes could also have contributed to an underwater hearing mechanism¹⁵. Associated with intracranial mobility and an open spiracle, the stapes in early tetrapods might thus have retained a function more directly

comparable with that of the fish hyomandibula, in that it controlled palatal and spiracular movements during breathing.

The close morphological similarities between the stapes of three otherwise distantly related tetrapods indicates that the massive stapes is a tetrapod autapomorphy, derived with respect to that of any non-tetrapod outgroup, but nonetheless primitive. Other forms, including the rod-like type found in tympanic ears, were later derivations. Alternatively, the massive stapes could be a synapomorphy uniting 'labyrinthodonts'³, not otherwise considered to be a natural group¹⁶. The relationship of *Acanthostega* to any 'labyrinthodont' is not yet supported by other synapomorphies. It has recently been suggested that possession of a double-headed stapes was the primitive tetrapod condition, comparable with the double-headed hyomandibula seen in osteolepiform fishes, and the stapes of many early reptiles¹⁷. This view would be corroborated by finds of double-headed stapes in the earliest tetrapods, but is not supported so far by the condition actually found in those where it is known. □

Received 10 March 1989; accepted 2 October 1989.

1. Watson, D. M. S. *Proc. zool. Soc. Lond.* B **1917**, 167–186 (1917).
2. Panchen, A. L. in *Studies in Vertebrate Evolution* (eds Joysey, K. A. & Kemp, T. S.) 65–68 (Oliver and Boyd, Edinburgh, 1972).
3. Bolt, J. R. & Lombard, R. E. *Biol. J. Linn. Soc.* **24**, 83–99 (1985).
4. Panchen, A. L. *Nature* **333**, 704 (1988).
5. Smithson, T. R. *Zool. J. Linn. Soc.* **76**, 29–90 (1982).
6. Clack, J. A. *Zool. J. Linn. Soc.* **79**, 121–148 (1983).
7. Carroll, R. L. in *The Terrestrial Environment and the Origin of Land Vertebrates* (ed. Panchen, A. L.) *Syst. Ass. Spec.* Vol. **15**, 293–317 (Academic, London, 1980).
8. Jaslow, A. P., Hetherington, T. E. & Lombard, R. E. in *The Evolution of the Amphibian Auditory System* (ed. Fritsch, B.) 69–91 (Wiley, New York, 1988).

9. Clack, J. A. *Phil. Trans. R. Soc. B* **318**, 1–107 (1987).
10. Clack, J. A. *Palaeontology* **31**, 699–724 (1988).
11. Randall, D. J., Burggren, W. W., Farrell, A. P. & Haswell, M. S. *The Evolution of Air-breathing in Vertebrates* (Cambridge University Press, 1981).
12. Allis, E. P. *J. Anat.* **56**, 189–294 (1922).
13. Panchen, A. L. *Phil. Trans. R. Soc. B* **263**, 279–326 (1972).
14. Carroll, R. L. in Panchen A. L. *Phil. Trans. R. Soc. B* **309**, 461–568 (1985).
15. Panchen, A. L. *Phil. Trans. R. Soc. B* **309**, 461–568 (1985).
16. Smithson, T. R. *Zool. J. Linn. Soc.* **85**, 317–410 (1985).
17. Lombard, R. E. & Bolt, J. R. in *The Evolution of the Amphibian Auditory System* (ed. Fritsch, B.) 37–67 (Wiley, New York, 1988).

Ocular dominance plasticity in adult cat visual cortex after transplantation of cultured astrocytes

Christian M. Müller & Johannes Best*

Max-Planck-Institut für Hirnforschung, Deutschordenstrasse 46, D-6000 Frankfurt Main 71, FRG

DURING a critical restricted period of postnatal development, the visual cortical circuitry is susceptible to modifications that are dependent on experience¹. If vision is restricted to only one eye during this period, the territories innervated by the deprived eye shrink considerably, whereas those innervated by the non-deprived eye expand, and the deprived eye loses the ability to influence almost all of the cells in the cortex^{2,3}. Thus, changes in ocular dominance are paralleled and possibly mediated by synapse elimination and axonal sprouting. Hypotheses about the mechanisms underlying ocular-dominance plasticity assume the activation of NMDA (N-methyl-D-aspartate) receptors and subsequent calcium influx as a trigger of synaptic modifications^{1,4,5}. In addition, plasticity relies on functional neuromodulatory afferents^{6,7}. On the basis of immunocytochemical studies, it was recently proposed that the presence of immature astrocytes is a prerequisite for visual cortical plasticity, and that the end of the critical period is causally linked to the maturation of astrocytes^{8,9}. Here we report, in support of this hypothesis, that resupplementation of the visual cortex of adult cats with astrocytes cultured from the visual cortex of newborn kittens reinduces ocular-dominance plasticity in adult animals.

The study described here was designed to test the hypothesis that the presence of immature astrocytes is a prerequisite for

ocular dominance (OD) plasticity by transplanting astrocytes cultured from newborn kittens into the visual cortices of five adult cats (Fig. 1). These animals were at least one year old and had therefore passed the critical period for cortical plasticity (for a review see ref. 1). Controls included animals ($N = 2$) that each received a transplantation of fibroblasts (the chief contaminant of the astrocytic cultures) in one hemisphere. In addition, every animal received a transplantation of cells destroyed by repeated freezing in the hemisphere contralateral to the one transplanted with intact cells. After 4–8 weeks of monocular deprivation, which began at the time of transplantation, animals were reanaesthetized and the OD distribution assessed using standard electrophysiological techniques. To avoid a bias caused by preferential recording within certain OD columns, electrode tracks were aimed in an oblique direction in the cortex. Recording tracks were placed at the lateral coordinates where the transplants had been performed. For each cell we determined the OD (five-class scale) and the response strength (four-class scale) by presenting optimally oriented, moving light-bars with a hand lamp on a tangent screen.

The mean response strengths observed in the experimental hemispheres (transplanted with intact cells) and the control hemispheres (transplanted with cells frozen before transplantation) do not differ significantly from one another (control: 2.98 ± 0.67 , $N = 321$; experimental: 3.00 ± 0.65 , $N = 548$). Significant differences in the OD distributions exist between experimental and control hemispheres, however. The OD distributions of the control hemispheres that received a transplantation of cells destroyed by multiple freezing showed a contralateral bias typical for normally raised adult cats^{10,11}. Most cells were driven by stimulation of either eye, and the OD distribution showed a moderate bias towards the contralateral eye (Fig. 2; far-left histogram, control). This bias occurred irrespective of whether the ipsilateral- or contralateral eye was closed during the period of monocular deprivation as a result of suturing (see Fig. 3 for data from individual animals). By contrast, the hemispheres that received intact astrocyte suspension showed OD distributions always biased towards the non deprived eye. In cases in which

* Present address: Ruhr-Universität Bochum, Institut für Neuroinformatik, 4630 Bochum 1, FRG

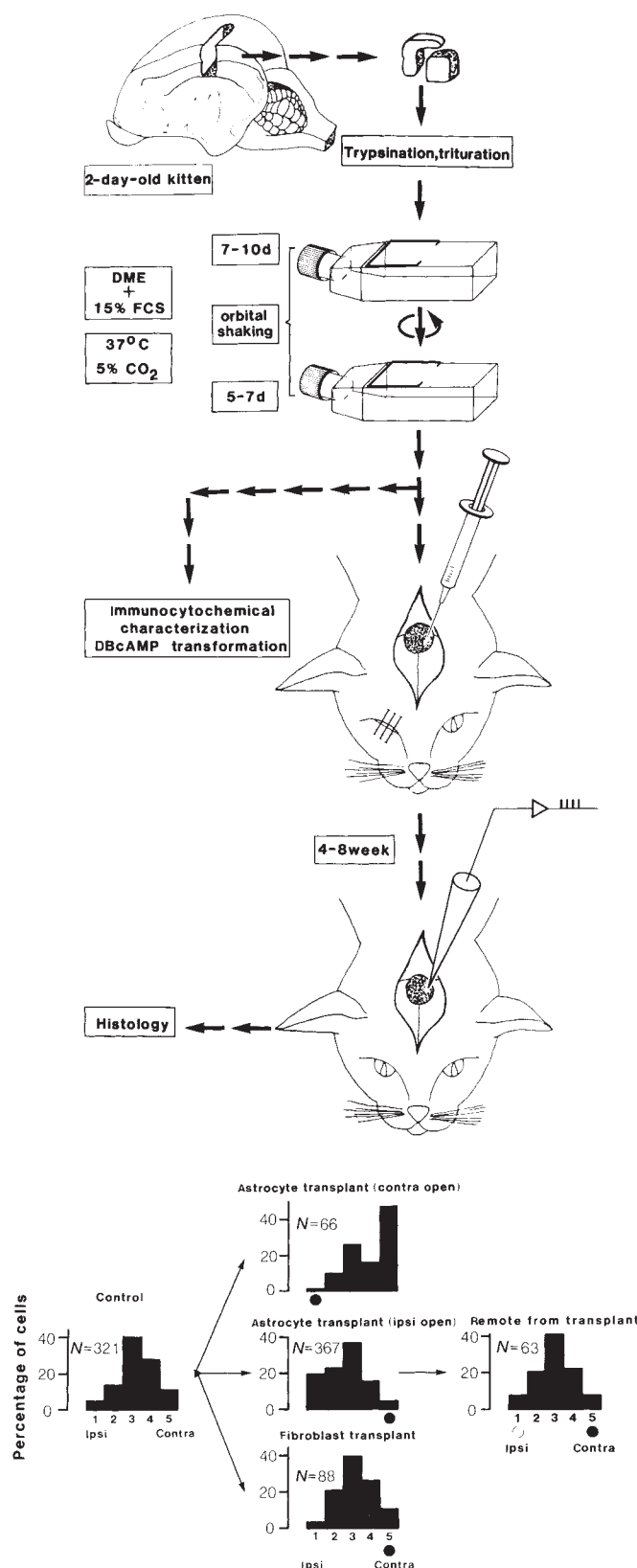


FIG. 1 Schematic representation of the experimental paradigm. Astrocytic cultures were prepared from the visual cortex of newborn kittens and injected as a cell suspension into the visual cortex of adult cats. At the same time one eye was sutured close. After a period of 4-8 weeks the OD distribution in the hemispheres was assessed by electrophysiological recording.

METHODS. Glial cultures were prepared from 2-day-old kittens. After decapitation, a block of visual cortex was removed, cleared of meningeal tissue and white matter, and minced in small pieces by use of two scalpel blades. After preincubation in Ca^{2+} - and Mg^{2+} -free phosphate buffer (CMF, 37 °C, 10 min), the tissue was transferred to a trypsin solution (1% in CMF, 37 °C, 10 min). After three washes with Dulbecco modified Eagles medium containing 10% fetal calf serum, the tissue was triturated with a fire-polished glass pipette (10 strokes) and seeded in precoated culture flasks (Primaria, 25 cm²) at a density of 30,000 cells cm⁻². The culturing medium (DMEM, 10-15% FCS) was changed every second day. Contaminating oligodendrocytes were removed by orbital shaking²⁸ at day 7-10. Fibroblast cultures were prepared from meningeal tissue under identical culturing conditions. When cultures had again reached confluency, cells were collected either by mild trypsinization or by mechanical scraping, centrifuged and resuspended in 40 μ l CMF. Part of the cultures were reseeded on glass coverslips, and the astrocytic nature of the cells was reconfirmed by immunostaining with the astroglial markers vimentin, S-100 and GFAP, in addition to testing the ability of dibutyryl cyclic AMP to transform the cells into process-bearing stellate cells. Quantification of the cultures revealed that 94% ($\pm$ 4%) of the cells were immunoreactive for both GFAP and vimentin and transformed into stellate cells after dibutyryl cAMP. The remaining cells were immunoreactive for vimentin only. Most of these cells had a flat morphology even after 24 h of treatment with dibutyryl cAMP and, thus, resembled fibroblasts. A few small cells containing vacuoles and resembling macrophages were encountered in astrocyte-enriched, as well as in fibroblast cultures. After collecting, the cell suspension was immediately injected into the visual cortex of anaesthetized adult cats mounted in a stereotaxic frame. One hemisphere received a total of 10-20 μ l (2-8 injection sites) of intact cell suspension, and the other was injected with the same amount of cell suspension frozen and thawed three times. Transplants were placed on the top of the gyrus suprasplenius at antero-posterior coordinates -5 to +2. After cell transplantation, one eyelid was sutured closed. The animals received antibiotic treatment for three days. After a period of 4-8 weeks of monocular deprivation, the animals were reanaesthetized, and the OD distribution in the cortices was assessed by standard electrophysiological techniques, using laquer-insulated tungsten electrodes. After termination of the experiments, animals were perfused with fixative, and frozen sections were prepared for reconstruction of penetration tracks, using small electrolytic lesions as references.

FIG. 2 OD histograms composed from data from seven cats. Monocular cells responding exclusively to stimulation of one eye are categorized in class 1 (driven by ipsilateral eye) or class 5 (driven by contralateral eye). Classes 2 and 4 represent binocular cells in which one eye was dominant. Class 3 represents cells responding equally well to stimulation of either eye. Histograms from the control hemispheres of all seven cats (far-left histogram, control) reveal a moderate bias towards the contralateral eye. The middle OD histograms were obtained from transplanted hemispheres. Combining astrocyte transplantation with deprivation of the ipsilateral eye resulted in an increased occurrence of cells driven exclusively by the experienced contralateral eye (class 5, top histogram, contra open). The central OD (ipsi open) was obtained from the astrocyte-transplanted hemisphere of four cats in which the contralateral eye was sutured closed. Note the bias towards the nondeprived ipsilateral eye. This effect was not present when recordings were obtained from locations remote from the transplantation site (far-right histogram, composite data from two animals). The bottom histogram shows data from two cats that received a transplantation of fibroblasts instead of astrocytes; the OD distribution does not differ from that obtained from control hemispheres. ○, Eye open; ●, eye closed.

the contralateral eye was closed, there is an ipsilateral dominance, whereas in one animal in which the ipsilateral eye was deprived, an overproportional number of cells were responsive only to contralateral stimulation (OD class 5; see Figs 2 and 3). The differences between the OD distributions of the control hemispheres and the astrocyte-receiving hemispheres are highly significant (χ^2 test, 4 degrees of freedom) in each cat that received an astrocyte transplantation (see Fig. 3 legend).

In addition to the control in which destroyed cells were injected into one hemisphere, two further controls were performed. First, the regional specificity of the effect of astrocyte transplantation was assessed by recording cells within the transplanted hemisphere of two animals, just 1 mm lateral to the transplant. This OD histogram does not differ significantly from the one obtained in the control hemisphere (Fig. 2). The local specificity was further detailed by separating cells recorded

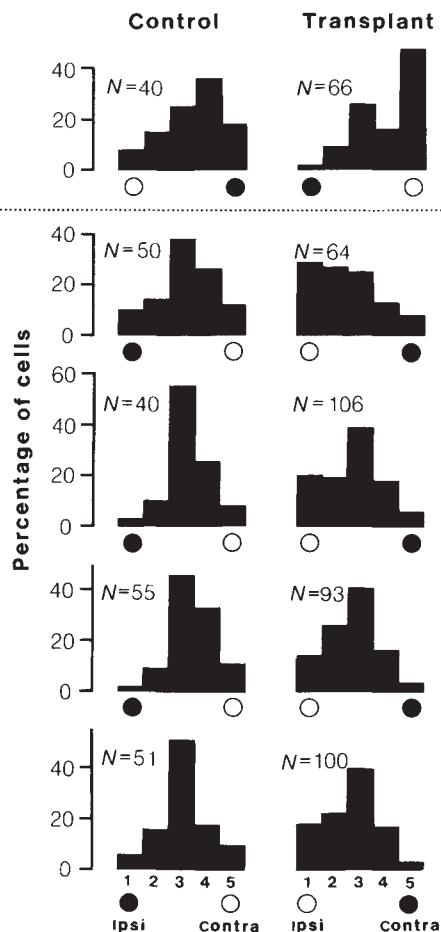


FIG. 3 OD histograms from visual cortex of all five monocularly deprived adult cats that received transplantation of cultured astrocytes. In the control hemispheres (left column), which received previously frozen cell suspension, OD histograms reflect a moderate bias towards the contralateral eye, irrespective of whether this eye was open or closed during the deprivation period (○, open; ●, closed). By contrast, in hemispheres that received intact astrocyte transplantation, OD histograms are always biased towards the open eye. This results in an overproportional presence of contralaterally dominated cells when the ipsilateral eye was sutured closed (row 1), and vice versa (rows 2–5). The overall differences in the OD histograms from the experimental and control hemispheres are statistically significant when tested with the χ^2 test. Calculated significances for each animal from top to bottom are: (1) $\chi^2=28.97$, $P<0.001$; (2) $\chi^2=20.60$, $P<0.001$; (3) $\chi^2=12.23$, $P<0.01$; (4) $\chi^2=31.52$, $P<0.001$; (5) $\chi^2=14.50$, $P<0.01$.

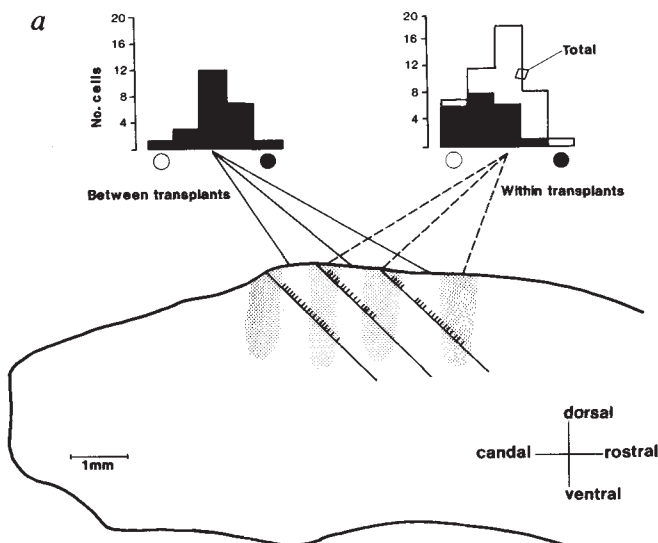
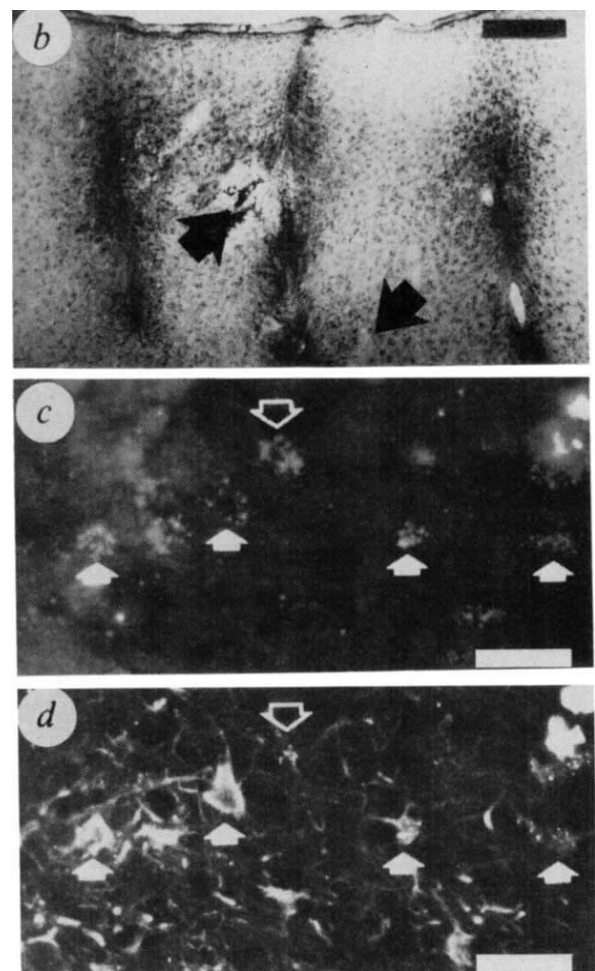


FIG. 4 *a*, OD histograms for cells recorded along the three penetrations passing through locations where transplanted astrocytes were present (dotted). Cells located on the border of the transplants ($N=7$) and two unresponsive units were omitted. OD histograms for cells recorded outside the transplants show a moderate bias towards the contralateral (deprived) eye, whereas a clear bias towards the (experienced) ipsilateral eye is seen in the OD distribution from cells recorded within transplanted regions. The composite histogram of the entire sample of cells given by the open columns in the right histogram continues to show a moderate bias towards the non-deprived eye. *b*, Low-power photomicrograph of the section drawn in *a*, stained with a monoclonal antibody directed against GFAP. Intense immunostaining is present at the sites of astrocytic transplants. The three transplantation sites correspond to rostral transplants in *a*. Two electrolytic lesions (arrows) mark an electrode track passing through the central transplantation site. *c*, *d*, Higher-power photomicrographs from the border region of an astrocyte transplant viewed with a fluorescent microscope. Transplanted cells were identified by the presence of rhodamine-labelled latex microspheres (*c*); the astrocytic nature of these cells was shown by the presence of GFAP-immunoreactivity (*d*). Note that most rhodamine-labelled cells also display GFAP immunoreactivity (for example, those marked by closed arrows in *c* and *d*). A subpopulation of cells containing rhodamine-labelled particles (open arrow in *c* and *d*) are not GFAP positive and thus may correspond to non-astrocytic cells such as fibroblasts or macrophages. Scale bars; *b*, 500 μm ; *c* and *d*, 25 μm .

METHODS. *a*, The location of the transplants was made visible by the presence of cells prelabelled with rhodamine-coupled microspheres fed to the astrocytic cultures (10 μl per 7 ml culture medium; 12 h) before transplantation (see *c*, *d*). Electrode tracks were reconstructed using electrolytic lesions as references (see *b*). *b*, Vibratome sections (50 μm) were incubated with a monoclonal antibody directed against GFAP (Boehringer). Antigen-locations were visualized using the Avidin-biotin method with diaminobenzidine as the chromogen⁸. *c*, *d*, For concurrent visualization of fluorescent latex microspheres, fed to the astrocytic cultures before transplantation, and GFAP, sections were incubated with the monoclonal antibody to GFAP made visible by a fluorescein isothiocyanate-conjugated antibody.



within and outside transplants. Electrode tracks were reconstructed with reference to small electrolytic lesions (Fig. 4b) and matched with transplantation sites made visible by pre-labelling the transplanted cells with fluorescent latex microspheres in two animals. Microscopic observation on sections stained with an antibody against glial fibrillary acidic protein (GFAP)⁸, an astrocytic marker, revealed that transplanted astrocytes were localized only in the vicinity of the transplants (Fig. 4). The cells were located in a region displaying an increased immunoreactivity for GFAP (Fig. 4b), which reflected the reactive gliosis due to local damage by the injection. Reactive gliosis was found also in regions where non-astrocytic cells and pre-frozen cells had been injected. The OD shift to the non-deprived (ipsilateral) eye was only present in the cell population situated within the transplants (Fig. 4a). In a second control experiment, we implanted two additional cats with cultured fibroblasts instead of astrocytes. The OD histogram from the transplanted hemispheres does not differ significantly from those obtained from the control hemispheres, which had received previously frozen cell suspension (Fig. 2).

These data show that the transplantation of astrocytes cultured from the cortices of newborn kittens is sufficient to produce OD changes induced by light deprivation in adult cats. The finding that injection of previously frozen cell suspensions and of cultured fibroblasts does not lead to OD shifts indicates that the transplantation of cells *per se*, or reactive gliosis, which is always associated with brain damage, is not sufficient to induce OD plasticity in adult cats. Reactive gliosis is mediated only to a minor extent by glial proliferation^{12,13}, and hence may not lead to the generation of a sufficient number of immature astrocytes. In addition, it seems unlikely that macrophages are the relevant cell type in the transplanted cultured cells because (1) the cells constituted only a minor fraction of the cultures (<2%), and (2) a high density of macrophages has to be expected in all transplanted regions because of their massive proliferation and invasion after local cortical damage¹⁴.

How do astrocytes support visual cortical plasticity, specifically axonal sprouting and synapse elimination? Several neurotrophic factors, such as nerve growth factor, basic fibroblast growth factor and the S-100 protein, are known to be synthesized and released by astrocytes^{15,16}. The ability to release trophic factors is developmentally regulated¹⁵. It is conceivable that such factors are released by immature astrocytes, and are selectively taken up by active terminals to support axonal sprouting or neurite extension. Furthermore, synapse elimination could be mediated by active separation of synapses by glial processes, and subsequent phagocytosis of the terminals. Such a mechanism has been observed in the normal development of the spinal cord¹⁷⁻¹⁹. A similar involvement of astrocytic processes in the regulation of synapses is found in the hypothalamo-neurohypophyseal system²⁰⁻²³. Recently it was shown that the increase of synapse density which occurs in glia-free cerebellar cultures is only reversed when astrocytes are transplanted to the cultures²⁴. The temporal limitation of synapse elimination to periods when immature astrocytes are present could be due to the loss of the astrocytes' ability during maturation to undergo morphological changes and to perform phagocytosis^{25,26}.

The data presented here further support the hypothesis that immature glial cells are causally involved in visual cortical plasticity. The mechanisms underlying the experience-dependent change of cortical connectivity discussed above could also explain the beneficial effect of transplanted astrocytes on behavioural recovery after cortical damage²⁷. It remains to be shown, however, whether astrocytes actively participate in the plastic changes proper. □

3. Swindale, N. V., Vital-Durand, F. & Blakemore, C. *Proc. R. Soc. B* **213**, 435-450 (1981).
4. Kleinschmidt, A., Bear, M. F. & Singer, W. *Science* **238**, 355-358 (1987).
5. Rauschecker, J. P. & Hahn, S. *Nature* **326**, 183-185 (1987).
6. Bear, M. F. & Singer, W. *Nature* **320**, 172-176 (1986).
7. Kasamatsu, T. & Pettigrew, J. D. *Science* **194**, 206-209 (1976).
8. Müller, C. M. *J. comp. Neurol.* (manuscript submitted).
9. Müller, C. M. *Glia* (manuscript submitted).
10. Hubel, D. H. & Wiesel, T. N. *J. Neurophysiol.* **28**, 1041-1059 (1965).
11. Levitt, F. B. & van Sluyters, R. C. *Dev. Brain Res.* **3**, 323-327 (1982).
12. Ferrara, N., Ousley, F. & Gospodarowicz, D. *Brain Res.* **462**, 223-232 (1988).
13. Furukawa, S., Furukawa, Y., Satoyoshi, E. & Hayashi, K. *Biochem. biophys. Res. Commun.* **136**, 57-63 (1986).
14. Miyake, T., Hattori, T., Fukuda, M. & Kitamura, T. *Brain Res.* **489**, 31-40 (1989).
15. Rudge, J. S., Manthorpe, M. & Varon, S. *Dev. Brain Res.* **19**, 161-172 (1985).
16. van Eldik, L. J. & Zimmer, D. B. *Brain Res.* **436**, 367-370 (1987).
17. Ronnevi, L. O. & Conradi, S. *Brain Res.* **80**, 335-339 (1974).
18. Ronnevi, L. O. *J. Neurocytol.* **6**, 487-504 (1977).
19. Ronnevi, L. O. *Cell Tiss. Res.* **189**, 203-217 (1978).
20. Tweedle, C. D. & Hatton, G. I. *Cell Tiss. Res.* **18**, 59-72 (1977).
21. Tweedle, C. D. & Hatton, G. I. *Brain Res.* **309**, 373-376 (1984).
22. Perlmuter, L. S., Tweedle, C. D. & Hatton, G. I. *Neuroscience* **12**, 503-511 (1984).
23. Perlmuter, L. S., Tweedle, C. D. & Hatton, G. I. *Neuroscience* **13**, 769-779 (1984).
24. Meshul, C. K., Seil, F. J. & Herndon, R. M. *Brain Res.* **402**, 139-145 (1987).
25. Skaper, S. D., Facci, L., Rudge, J., Katoh-Semba, R., Manthorpe, M. & Varon, S. *Dev. Brain Res.* **25**, 21-31 (1986).
26. Saad, Y. A., Al-Zuhair, A. G. H. & Dawod, B. *Glia* **1**, 211-218 (1988).
27. Kesslak, J. P., Nieto-Sampedro, M., Globus, J. & Cotman, C. W. *Expl. Neurol.* **92**, 377-390 (1986).
28. McCarthy, K. D. & de Vellis, J. *J. Cell Biol.* **85**, 890-902 (1980).

ACKNOWLEDGEMENTS. We are indebted to Dr H. D. Hofmann and Susanne Wallenstein for support in establishing the tissue cultures, and to Nancy Gray for improving the language.

Target size regulates calibre and myelination of sympathetic axons

James T. Voyvodic*

Department of Anatomy and Neurobiology,
Washington University Medical School, 660 S. Euclid, St Louis,
Missouri 63110, USA

AXONS in vertebrate peripheral nerves are ensheathed by Schwann cells. For some axons, this sheath consists of a single layer of glial cell cytoplasm and plasma membranes; for other axons, Schwann cells form multilayered myelin. Whether or not a Schwann cell makes myelin is determined by a signal from the axon^{1,2}, but the nature of this signal is not known. Here I show that sympathetic postganglionic axons, which are normally not myelinated, become myelinated when their calibre is increased as a result of increasing the size of the peripheral target they innervate. This result implies that axon calibre, which is known to be correlated with myelination³, is in fact the crucial determinant of whether an axon becomes myelinated. Furthermore, the finding that increasing or decreasing target size causes corresponding increases or decreases in axon size indicates that axon calibre is itself regulated by retrograde signals from peripheral target tissues.

To study the effect of peripheral target size on ganglion cell axons, the relative size of a peripheral target of the superior cervical ganglion was altered during development by changing the ratio between the amount of target tissue and the number of innervating nerve cells in the ganglion. The target tissue tested in these experiments was the rat submandibular salivary gland. To make a smaller target than normal, the submandibular duct was ligated in 4-week-old rats; this results in glands weighing approximately 20% of normal when the animals reach 30 weeks of age⁴⁻⁶. These smaller glands were innervated by a normal number of superior cervical ganglion neurons (3,800 axons on average). In a second group of rats, relative target size was increased by cutting one branch of the sympathetic nerve innervating the gland at birth. When these animals were examined at 30 weeks of age, the partially denervated glands were of normal size, but were innervated by only 5-50% the normal number of sympathetic neurons⁶.

Received 20 July; accepted 2 October 1989.

1. Singer, W. in *The Neural and Molecular Bases of Learning* (eds Changeux, J. P. & Konishi, M.) 301-336 (Wiley, Chichester, 1987).

2. Le Vay, S., Stryker, M. P. & Shatz, C. J. *J. comp. Neurol.* **179**, 223-244 (1978).

* Present address: Department of Biology (Medawar Building), University College London, London WC1E 6BT, UK.

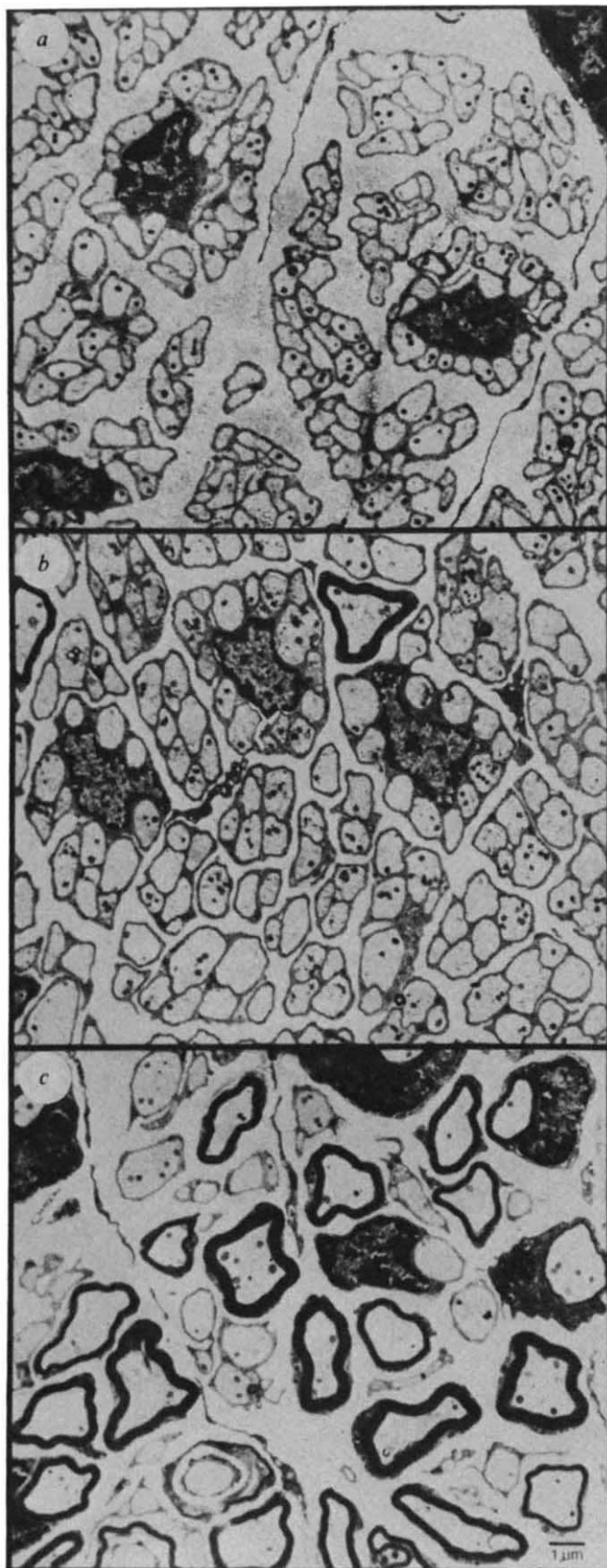


FIG. 1 Effect of target size on the calibre of postganglionic sympathetic axons innervating the submandibular gland. Axons innervating: *a*, small target (submandibular duct ligated at 4 weeks of age); *b*, control target; *c*, large target (gland partially denervated at birth). All determinations were made in adult animals (30 weeks). Mean axon diameters were: *a*, $0.80 \pm 0.01 \mu\text{m}$; *b*, $1.15 \pm 0.01 \mu\text{m}$; *c*, $1.43 \pm 0.02 \mu\text{m}$. A portion of nerve midway between the gland and the external carotid nerve trunk was removed from 6-month-old animals after death; nerves were fixed by immersion and processed for electron microscopy as previously described⁶. Axon circumference was measured from electron micrographs (4,000 $\times$) using a digitizing tablet, and axon diameters were calculated as circumference/ π . Scale bar is the same for all 3 panels.

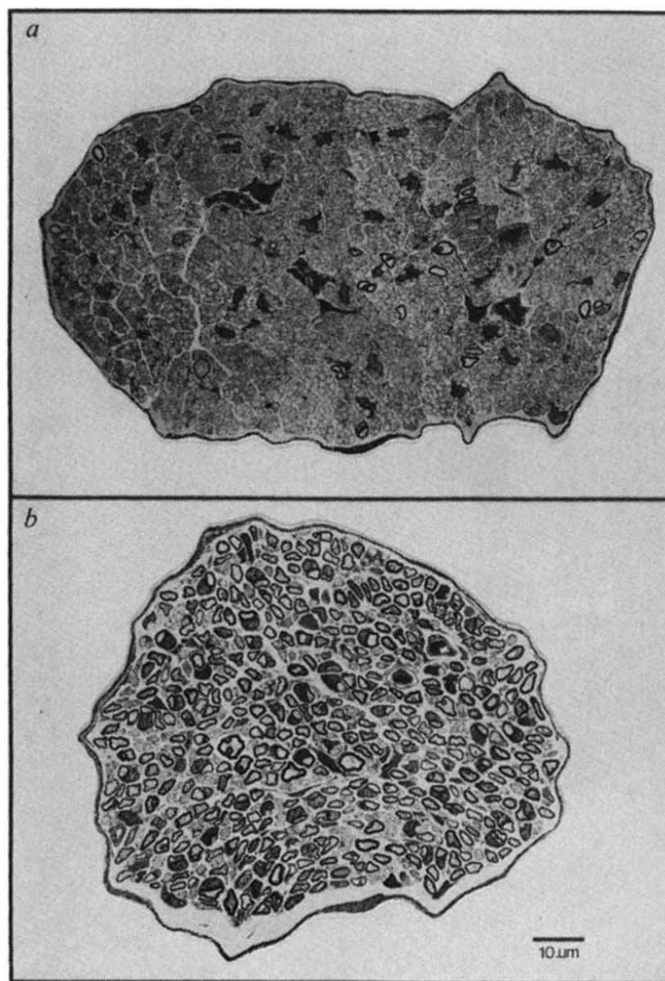


FIG. 2 Effect of increasing target size on the myelination of sympathetic axons. Electron microscope photomontages of whole postganglionic nerves innervating the submandibular gland in *a*, control (sham operated) nerve and *b*, nerve innervating relatively enlarged target (partially denervated at birth); both nerves are shown at the same magnification. In *a*, there are 3,192 unmyelinated and 25 myelinated axons; in *b*, 304 axons are unmyelinated and 386 are myelinated. Nerves were prepared for electron microscopy as described in the legend to Fig. 1. Retrograde labelling of the postganglionic nerve with horseradish peroxidase confirmed that at least 95% of the axons seen in cross-sections were axons of superior cervical ganglion neurons⁶.

The postganglionic nerves carrying sympathetic axons to the submandibular gland were cut in cross-section and examined by electron microscopy (Fig. 1). Decreasing relative target size resulted in axons whose diameters were on average 31% ($\pm 0.5\%$ s.e.m.) smaller than control axons, whereas increasing relative

target size resulted in axons whose diameters were 24% ($\pm 1.7\%$) larger than normal. These changes in axon diameter are comparable to the magnitudes of changes previously observed in cell soma diameter and dendritic length when target size is altered under similar experimental conditions⁶. Taken together, these observations suggest that retrograde signals from the target provide a general mechanism for regulating the size of superior cervical ganglion neurons and their processes.

Target-induced changes in the size of the axon were associated with significant changes in the way in which these axons were

TABLE 1 Effect of axonal target size and preganglionic denervation on Schwann cells associated with postganglionic sympathetic axons

Experimental condition	Number of animals	Target size ($\mu\text{g}/\text{axon}$)	Total number of axons	Number of myelinated axons	Total number of Schwann cells/cross-section
Small target	3	24 ± 4	$3,596 \pm 102$	13 ± 1	323 ± 19
Control	4	90 ± 4	$3,777 \pm 90$	24 ± 4	426 ± 39
Large target A	3	252 ± 56	$1,620 \pm 396$	440 ± 189	$1,111 \pm 352$
Large target B	2	532 ± 27	636 ± 38	291 ± 67	470 ± 75
Large target C	2	$1,463 \pm 155$	233 ± 15	140 ± 1	233 ± 15
Preganglionic denervation					
Control target	2	73 ± 3	$3,971 \pm 458$	19 ± 5	401 ± 22
Large target	2	954 ± 506	595 ± 292	141 ± 57	—

In the top part of the table, animals with relatively large targets were divided into 3 groups (A–C) because of the wide range of sizes produced by partial denervation (see text). Data in the lower part of the table show the effect of partial denervation of the submandibular gland for animals in which the preganglionic nerve was cut at birth⁶. The number of Schwann cells in a cross section of the nerve was estimated by dividing the total number of axons in the nerve by the number of axons per Schwann cell (calculated as described in Fig. 3.) Values are given $\pm$ s.e.m.

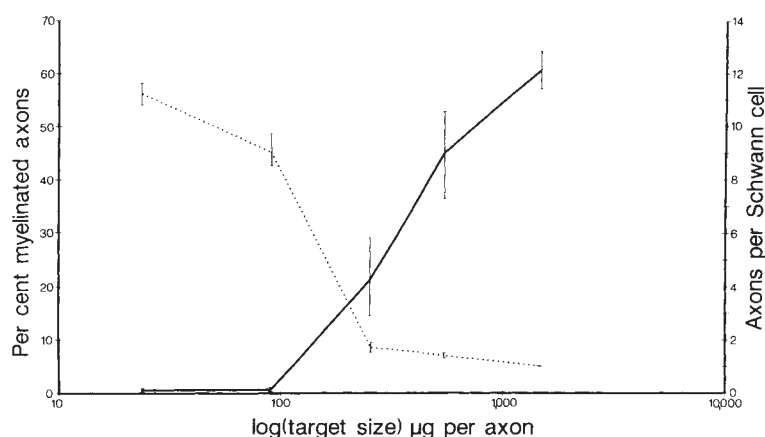


FIG. 3 Influence of target size on myelination, expressed as the percentage of axons that are myelinated (unbroken line), and on the number of axons ensheathed by each Schwann cell (dotted line). Target size is defined as the mass of the submandibular gland divided by the number of innervating sympathetic axons (counted in photomontages of whole nerves, as in Fig. 2). Each data point is the mean of values from several animals (Table 1). The number of axons ensheathed per Schwann cell was counted for each glial cell whose nucleus was visible in the plane of section; axon profiles in direct contact with the Schwann cell membrane were considered to be ensheathed by that glial cell.

ensheathed by Schwann cells. In control animals, 99.3% ($\pm 0.1\%$ s.e.m.) of the axons in the postganglionic nerve were unmyelinated. When relative target size was increased by partial denervation, there was a dramatic increase in the number of myelinated axons in the nerve (Fig. 2). The absolute number of myelinated axons increased twelve times, on average, over the number in control nerves. Moreover, the fraction of axons that was myelinated increased progressively with increasing size of the peripheral target; for the largest targets, 60% of the axons were myelinated (Table 1; Fig. 3). When postganglionic nerves innervating experimentally enlarged targets were examined at multiple sites between the ganglion and the gland, the number of myelinated axons counted at each level was significantly above the number present in controls, indicating that these axons were myelinated along their entire length. Finally, an increase in myelination could be induced by increasing target size even if the superior cervical ganglion was denervated at birth (Table 1), suggesting that the increased myelination does not depend on electrical activity in the ganglion cells.

Two lines of evidence in this study indicate that the signal for Schwann cells to myelinate sympathetic axons is an increase in axon calibre. First, in agreement with previous reports^{3,7,8}, there was a clear size threshold above which axons became myelinated. Whereas 95% of unmyelinated axons were less than $1.6 \mu\text{m}$ in diameter, regardless of the experimental conditions, 90% of myelinated axons were greater than $1.6 \mu\text{m}$ in diameter. Second, even for axons below this size threshold for myelination, the number of axons ensheathed by each Schwann cell varied systematically with target size and axon calibre. Thus, when axon calibre was decreased by making a smaller target, there was no significant change in the number of myelinated axons in the postganglionic nerve, but there was a significant increase in the average number of axons ensheathed by each glial cell. On the other hand, as target size and axon calibre were increased in these experiments, the average number of axons associated with each glial cell gradually decreased (Fig. 3). This decrease

was accompanied by an increase in the total number of Schwann cells present in the nerve (Table 1).

These systematic changes in the way in which Schwann cells ensheath axons of different sizes imply that Schwann cells are capable of responding to changes in axon calibre long before the axons reach the size threshold for myelination. Interestingly, the point at which Schwann cells switched from simple ensheathment to myelination was the point at which the ratio of axons to Schwann cells approached unity (Fig. 3). This relationship suggests that a Schwann cell that does not produce myelin might be triggered to begin myelination not because it detects some critical axon diameter, but rather because it reaches a point where it can no longer accommodate further axon growth by reducing the number of axons ensheathed. After Schwann cells switch to a myelinating phenotype, they once again respond to axon growth in a graded manner. In this case, however, the response is not to relinquish axons, but to add additional layers of compact myelin⁷.

These results imply that there are no qualitative differences between axons that normally become myelinated and those that do not: increasing the size of a normally unmyelinated axon is sufficient to induce Schwann cells to myelinate it. These experiments also show that the calibre of peripheral axons is regulated by the size of their target. This regulation provides a mechanism by which the nervous system can continuously adjust to the growth of target tissues by modifying axonal conduction properties through retrograde signals from the target; these signals affect both axon size and myelination. □

Received 18 August; accepted 4 October 1989.

- Weinberg, H. J. & Spencer, P. S. *Brain Res.* **133**, 363–378 (1975).
- Aguiayo, A. J. et al. *J. Neurocytol.* **5**, 565–573 (1976).
- Duncan, D. J. *comp. Neurol.* **60**, 437–472 (1934).
- Tamarin, A. in *Secretory Mechanisms of Salivary Glands* (ed. Schneyer, L. H. & Schneyer, C. A.) 220–237 (Academic, New York, 1966).
- Womble, M. D. & Roper, S. *J. Neurophysiol.* **58**, 276–287 (1987).

6. Voyvodic, J. T. *J. Neurosci.* **9**, 1997–2010 (1989).
 7. Sanders, F. K. *Proc. R. Soc. B* **135**, 323–357 (1948).
 8. Friede, R. & Samorajski, T. *J. comp. Neurol.* **130**, 223–231 (1967).

ACKNOWLEDGEMENTS. I am grateful to Dale Purves for advice and support throughout this study and to P. Newton and D. Dill for their invaluable technical assistance. I thank A. Wilson and T. Ramey for help in counting axons, and M. Raff and E. Rubin for useful comments on the manuscript. This work was supported by a NSF predoctoral fellowship and by NIH grants to D. Purves.

Membrane depolarization evokes neurotransmitter release in the absence of calcium entry

B. Hochner, H. Parnas & I. Parnas

The Otto Loewi Center, Department of Neurobiology,
The Hebrew University, Jerusalem, Israel

THE discovery that Ca^{2+} is necessary for the release of neurotransmitter, the primary means by which nerve cells communicate, led to the calcium hypothesis of neurotransmitter release^{1–4}, in which release is initiated after an action potential only by an increase in intracellular Ca^{2+} concentration near the release sites and is terminated (1–2 ms) by the rapid removal of Ca^{2+} . Since then, the calcium-voltage hypothesis has been proposed^{5,6}, in which the depolarization of the presynaptic terminals has two functions. First, in common with the calcium hypothesis, the Ca^{2+} conductance is increased, thereby permitting Ca^{2+} entry. Second, a conformational change is induced in a membrane molecule that renders it sensitive to Ca^{2+} , and then binding of Ca^{2+} to this active form triggers release of neurotransmitter. When the membrane is repolarized, the molecule is inactivated and release is terminated, regardless of the local Ca^{2+} concentration at that moment. This hypothesis, in contrast to the calcium hypothesis, accounts for the insensitivity of the time course of release to experimental manipulations of intracellular Ca^{2+} concentration^{7–11}. Furthermore, it explains rapid termination of release after depolarization, even though Ca^{2+} concentration may still be high. Here we describe experiments that distinguish between these two hypotheses and find that our results support the calcium voltage hypothesis.

A critical test to distinguish the two hypotheses for the mechanism of neurotransmitter release separate the two effects of depolarization by raising intracellular Ca^{2+} and inducing membrane depolarization without entry of Ca^{2+} . The calcium hypothesis would then predict that the membrane depolarization would not increase release, whereas the calcium-voltage hypothesis would predict an increase in release, locked in time to the depolarizing pulse. The introduction of photosensitive Ca^{2+} chelators, the 'caged Ca^{2+} ' compounds¹², makes such tests feasible. These compounds have a high affinity for Ca^{2+} , but after exposure to light, the affinity is reduced and Ca^{2+} is released. It is therefore possible to induce a phasic increase of Ca^{2+} concentration inside a presynaptic terminal without changing the membrane potential.

Zucker and Haydon¹³ used a caged Ca^{2+} compound, nitr-5, to test the calcium-voltage hypothesis, and found that asynchronous transmitter release was induced by photolysis of nitr-5. In the absence of Ca^{2+} entry, as depolarization did not affect this asynchronous release, it was concluded that membrane depolarization has no direct role in the release process. The experiments underlying this conclusion, however, were made on a synapse in culture, where release is slow. It is possible that, unlike in fast synapses, a change in membrane potential does not play a part in determining the time course of release, and/or that the experimental protocol was not designed to bring out such effects (see our delay histogram, Fig. 3). We used nitr-5 to test the calcium-voltage hypothesis at a conventional fast synapse, the neuromuscular junction of the opener muscle in the first walking leg of the small crayfish *Procambarus clarkii*. It has been indicated that in this preparation membrane potential

has a role in addition to the classical one of opening Ca^{2+} channels^{5,14}.

Figure 1 depicts the time course of the effects of pressure-injecting nitr-5 into the excitatory axon through an intracellular microelectrode placed close to the terminal¹⁵. Baseline responses to twin impulses (20 ms apart) recorded shortly after the intracellular penetration are shown in Fig. 1a. A reduction of release was apparent within three minutes after the start of injection (Fig. 1b). As time elapsed and more nitr-5 was injected into the axon, release was further reduced (Fig. 1c) until, after 18 min, it was nearly abolished (Fig. 1d). This reduction in release is expected because nitr-5 chelates Ca^{2+} and its K_D in darkness is close to the resting level of intracellular Ca^{2+} ($\sim 0.1 \mu\text{M}$). These results confirm that nitr-5 is reaching the nerve terminals. At this state, the frequency of stimulation was raised to load the terminal with Ca^{2+} and to allow release to be measurable (Fig. 2a). A flash of light was then delivered which resulted in a transient threefold elevation of the evoked release (Fig. 2b). About one second after the flash, release is stabilized at a level higher than the control (compare Fig. 2a with Fig. 2b), indicating that the increase of intracellular Ca^{2+} concentration achieved by photolysis of nitr-5 is in a range that is effective for release.

Given these results, nitr-5 could be directly used to test the unique feature of the calcium-voltage hypothesis. To this end, we blocked Ca^{2+} entry using an extracellular solution essentially devoid of Ca^{2+} . Addition of EGTA damages the preparation, so we added a high concentration of Mg^{2+} (12.5 mM) both to block any possible entry of Ca^{2+} and also to substitute for Ca^{2+} in opening postsynaptic glutamate channels¹⁶. The effect of

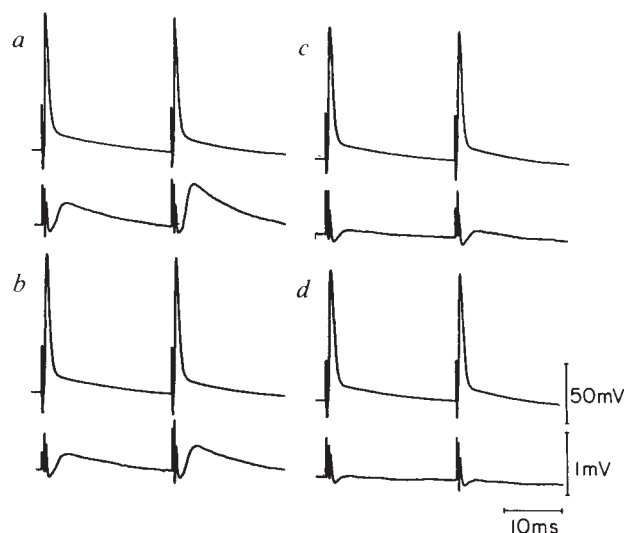


FIG. 1 a–d, Pressure-injection of nitr-5 reduces synaptic release. Upper traces show the axon action potential. Lower traces show the synaptic potentials. Twin impulses, 20 ms apart, were given at 3.3 Hz. Each record is an average of 50 sweeps. a, Control, before nitr-5 injection; and b, 3 min, c, 10 min and d, 18 min after beginning of injection of nitr-5. The excitatory axon was impaled distally to the main bifurcation¹⁵ by an intracellular microelectrode, which served for recording and for pressure injection of nitr-5. The axon was stimulated by a suction electrode at its entry to the opener muscle. Another intracellular electrode recorded synaptic potentials from a muscle fibre between these two sites.

METHODS. Crayfish were supplied by Atchafalaya, Louisiana. The first walking legs of small (body length 4–4.5 cm) animals were used. The bevelled microelectrode contained 50 mM nitr-5 (tetrasodium salt; Calbiochem), 35 mM CaCl_2 , 85 mM KCl, 10 mM Tris-HCl, pH 7.4, and 0.5% fast green. For pressure injection (1 s, 10–40 psi) the Pico-Injector Model PLI-100 (Medical System) was used. Between 40–100 injection pulses were given during a period of ~ 30 min. E.p.s.ps were recorded with a microelectrode containing 2.5 M potassium acetate (6–10 M Ω). The preparation was perfused with a solution containing: 220 mM NaCl, 5.4 mM KCl, 12 mM CaCl_2 , 2.5 mM MgCl_2 , 10 mM Tris-maleate, pH 7.6, at 18–20 °C. Transilluminating light was filtered to remove wavelengths below 475 nm to avoid nitr-5 photolysis.

voltage on the probability of release was assessed by constructing synaptic delay histograms^{3,5}. After the injection of nitr-5 and perfusion with the Ca^{2+} -entry blocking solution, evoked release was abolished. That is, miniature endplate potentials (m.e.p.ps) were about equally distributed 10 ms before and 10 ms after stimulus pulses (Fig. 3b). Then, while the nerve was being stimulated, three flashes of light (about 40 s apart) were given. The results were pooled in a histogram (Fig. 3c). Two changes occurred (compare Fig. 3b and c). First, the baseline level of spontaneous m.e.p.ps increased by more than 6 times. Second, an additional increase in the rate of release was observed in the time range 2–4 ms after the stimulus. That is, there was now evoked release. Furthermore, the time range of this evoked release was consistent with the time course of release in the presence of extracellular Ca^{2+} (refs 5, 14). We obtained similar results in four further experiments.

The increase in transmitter release after each flash was transient. For example, delay histograms constructed from the first 200 impulses (10 s) following the first flash contained 7 m.e.p.ps (in 10 ms following the stimulus), whereas the next 200 impulses resulted in only 2 m.e.p.ps. The transient increase in release indicates that after each flash the intracellular Ca^{2+} concentration increased briefly and that the Ca^{2+} was then removed by the intracellular Ca^{2+} -buffering systems. The observed evoked release could have been due to a temporary elevation of intracellular Ca^{2+} concentration. Alternatively, it could have resulted from a reduced ability of the photolysed

nitr-5 to chelate small amounts of Ca^{2+} that might have entered the terminal. As the reduction in the affinity of the photolysed nitr-5 is not reversible but the increased release was transient, we favour the first possibility.

The level of stimulus-evoked release in our experiments was rather lower than that seen in the presence of extracellular Ca^{2+} . It is possible that intracellular Ca^{2+} buffers compete with the nitr-5 to strip away some of its Ca^{2+} , diminishing the increase in Ca^{2+} after a flash. We therefore examined the effect of interfering with intracellular Ca^{2+} -buffering mechanisms using the uncoupler carbonyl cyanide *m*-chlorophenylhydrazone (CCCP)¹⁷. Two additional modifications were made in this experiment. First, 0.2 mM Ca^{2+} was added to the bathing medium to improve the ability of the axon to sustain action potentials. Second, as there was a small amount of Ca^{2+} in the medium, the measures for blocking Ca^{2+} entry were strengthened by including 2 mM Mn^{2+} in addition to the 12.5 mM Mg^{2+} . In the controls without nitr-5 injection, release could not be evoked at stimulation rates of 10 Hz. In 1,000 pulses, only six quanta were seen. This was so even in control trials where 3,4-diaminopyridine (0.1 mM) was added and the nerve action potential was clearly broadened. Figure 4a (traces 1 and 2, fast and slow time base, respectively) shows synaptic responses in the presence of extracellular Ca^{2+} 75 minutes after injection of nitr-5. Figure 4b shows the block of release in the Ca^{2+} -blocking medium. The responses seen in Fig. 4c were obtained 11 minutes after application of 10 μM CCCP. Two effects are evident. First, there was increase in the spontaneous release (Fig. 4c, trace 2), presumably resulting from an increase in intracellular Ca^{2+} concentration. The more striking result, however, was the appearance of synaptic potentials. These are seen even after single impulses without averaging. Now after a flash of light (Fig. 4d), the amplitude of evoked excitatory postsynaptic potentials (e.p.s.ps) and the frequency of spontaneous m.e.p.ps increased even more.

Interestingly, Zucker and Lando¹⁷, who applied 10 μM CCCP in the same preparation, did not observe evoked release. The difference between their results and ours may be explained by

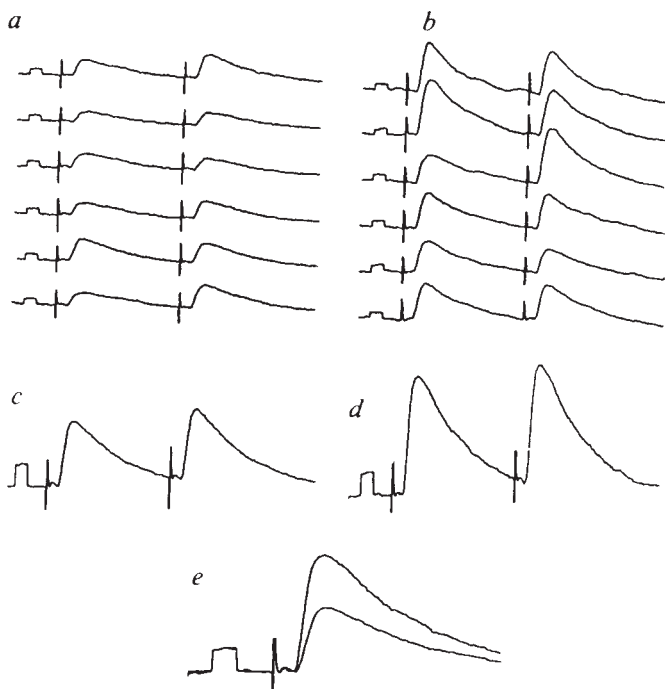


FIG. 2 Photolysis of nitr-5 caused an increase in evoked release. This is a continuation of the experiment shown in Fig. 1. Twin-impulse stimulation (20 ms intervals) at 10 Hz slowly restored release (compare with Fig. 1d). a, Single responses (sampled at 5 Hz) after e.p.s.ps reached a steady-state level. b, Single responses after a flash of light (top to bottom). c, An average of 10 sweeps before the light flash. d, An average of 5 sweeps after the light flash. e, Superposition of the first responses in c and d. Calibration pulse: 1 mV; 2 ms.

METHODS. The xenon lamp was an ensemble of a Strobex Power Pack, model 238 and a Strobex Lamp model 278 (Chadwick-Helmuth, California). The flash lamp was mounted in one of two configurations. In the first, the lamp was aimed with an angle of 45°, in the second it was positioned vertically over the preparation. The light was filtered to remove short-wave ultraviolet (<305 nm). The energy of the flash was about 180 joules for the first configuration and 100 joules for the second. A light flash without injection of nitr-5 had no effect. The light was focused with an elliptical mirror (Pichel, California) to give an image of ~7 mm at the focal plane.

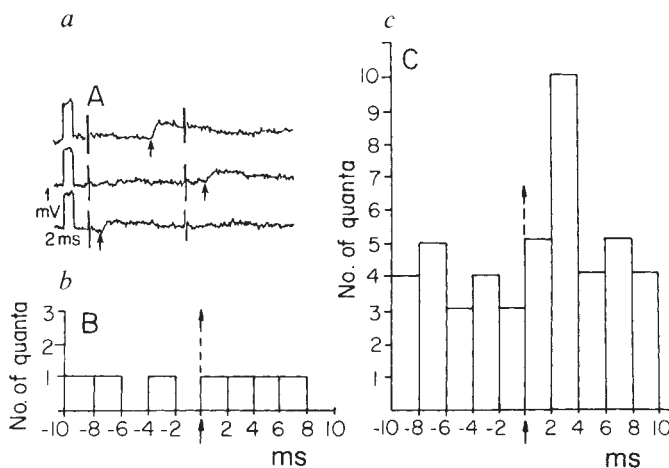


FIG. 3 Photolysis of nitr-5 induces evoked release under conditions of zero Ca^{2+} concentration and 12.5 mM Mg^{2+} in the external medium. Twin impulses (20 ms apart) were given at a repetition rate of 10 Hz. The delay from the stimulus to a quantum response was measured (arrows in a). Delay histograms (bin size 2 ms) describing the probability of occurrence of single events are given for 10 ms before the stimulus (negative timescale) and after the stimulus. Responses were measured for a period of 10 ms after the two impulses in a sweep and lumped together. For responses appearing before stimulation, measurements were taken for two successive periods of 10 ms before the first stimulus and lumped together. a, Examples of records after a flash of light. b, Delay histogram after injection of nitr-5 in the dark. c, Delay histogram in a nitr-5 loaded preparation following 3 consecutive flashes of light. For each histogram, 600 sweeps were taken; the histogram contains responses to 1,200 impulses. Arrows in b and c indicate point of stimulation.

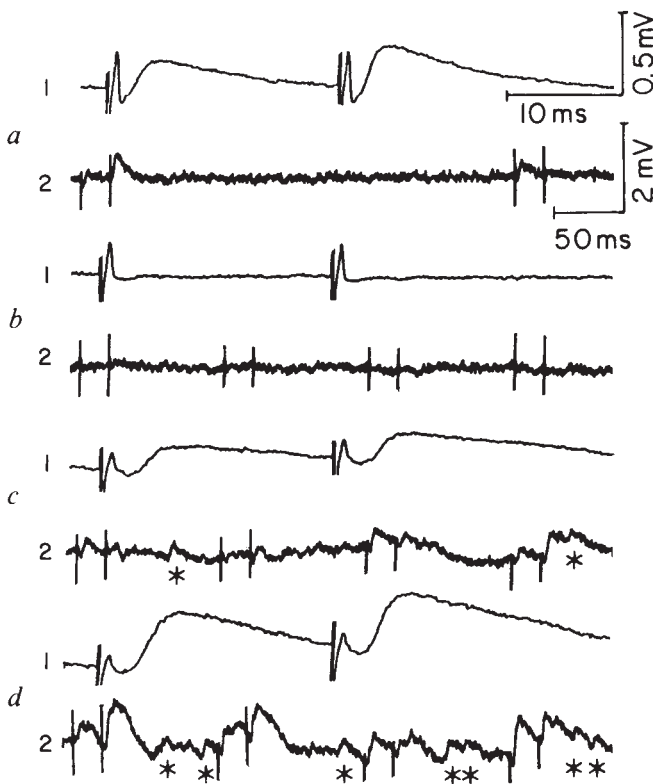


FIG. 4 Ca^{2+} loading of the nerve terminal supports action potential evoked release in the absence of Ca^{2+} entry. *a*, Trace 1 is an average of 100 traces taken at 3.3 Hz 75 min after the injection of nitr-5; trace 2 is a single trace, at a slow time base, showing evoked release but not spontaneous release. *b*, Records taken 4 min after wash with a solution containing 0.2 mM Ca^{2+} , 12.5 mM Mg^{2+} and 2 mM Mn^{2+} . The wash was very rapid and effective as the washing solution was directly aimed at the small space of the opener muscle in the shell. It is shown in trace 1 that evoked release is completely blocked, even at 10 Hz stimulation rate and in trace 2 that no m.e.p.s are seen. Eleven minutes after addition of 10 μM CCCP, evoked release can be seen in the average record in *c* (trace 1) and m.e.p.s (asterisks) in trace 2. The single trace (2) in *c* also shows evoked e.p.s.p.s. *d*, After a flash of light, a further increase in the size of the evoked response is evident in trace 1 (average of 50 traces), with a concomitant increase in m.e.p.s frequency in trace 2. In the averaged responses in *a*(1)–*d*(1), the nerve action potential appearing before the synaptic potential originates from cross talk between the two microelectrodes.

two factors: (1) their measurements were taken too long after the application of CCCP, and (2) the injection of nitr-5 might have postponed the build-up of a toxic level of Ca^{2+} concentration because of its Ca^{2+} -chelating properties. It is known that at a high intracellular Ca^{2+} concentration, evoked release is blocked while spontaneous release continues¹⁸. Indeed, we measured release shortly after application of CCCP, when the frequency of spontaneous m.e.p.s was still low.

We conclude that (1) after a flash of light in nitr-5-injected axons, particularly in the presence of CCCP, spontaneous release is enhanced, these m.e.p.s appearing randomly in time; (2) under conditions of raised intracellular Ca^{2+} , depolarizing impulses evoked synchronous release, despite the blocking of Ca^{2+} entry; and (3) the time course of this evoked release was normal. Therefore, in the fast synapse of the crayfish neuromuscular junction, membrane depolarization plays a part in the release process, in addition to its well-known effect of opening Ca^{2+} channels. □

Received 14 June; accepted 16 October 1989.

1. Katz, B. & Miledi, R. *Proc. Roy. Soc. B* **161**, 496–503 (1965).
2. Miledi, R. *Proc. Roy. Soc. B* **183**, 421–425 (1973).
3. Katz, B. & Miledi, R. *Proc. Roy. Soc. B* **161**, 483–495 (1965).
4. Smith, S. U. & Augustine, G. J. *Trends Neurosci.* **11**, 458–464 (1988).
5. Parnas, I., Parnas, H. & Dudel, J. *Pflügers Arch. Eur. J. Physiol.* **406**, 131–137 (1986).

6. Parnas, H., Parnas, I. & Segel, L. A. *Int. Rev. Neurobiol.* (in the press).
7. Barrett, E. F. & Stevens, C. F. *J. Physiol. Lond.* **227**, 691–708 (1972).
8. Matzner, H., Parnas, H. & Parnas, I. *J. Physiol. Lond.* **398**, 109–121 (1988).
9. Parnas, H., Hovav, G. & Parnas, I. *Biophys. J.* **55**, 859–874 (1989).
10. Dattner, N. B. & Gage, P. W. *J. Physiol. Lond.* **303**, 299–314 (1989).
11. Andreu, R. & Barrett, E. F. *J. Physiol. Lond.* **308**, 79–97 (1980).
12. Adams, S. R., Kao, J. P. Y., Grynkiewicz, G., Minta, A. & Tsien, R. Y. *J. Am. chem. Soc.* **110**, 3212–3220 (1988).
13. Zucker, R. S. & Haydon, P. G. *Nature* **335**, 360–362 (1988).
14. Parnas, H., Dudel, J. & Parnas, I. *Pflügers Arch. Eur. J. Physiol.* **496**, 121–130 (1986).
15. Wojtowicz, J. M. & Atwood, H. L. *J. Neurobiol.* **14**, 385–390 (1983).
16. Hatt, H., Franke, C. H. & Dudel, J. *Pflügers Arch. Eur. J. Physiol.* **411**, 17–26 (1988).
17. Zucker, R. S. & Lando, L. *Science* **231**, 574–579 (1986).
18. Adams, O. J., Takeda, K. & Umbach, J. A. *J. Physiol. Lond.* **369**, 145–159 (1985).

ACKNOWLEDGEMENTS. I. Parnas is the Greenfield Professor of Neurobiology. We thank Professor M. Devor for a critical reading of the manuscript. This research was supported by the Goldie Anna charitable trust.

Germ-line transmission of a disrupted β_2 -microglobulin gene produced by homologous recombination in embryonic stem cells

Maarten Zijlstra*, En Li*, Fereydoun Sajjadi†, Suresh Subramani† & Rudolf Jaenisch*

* Whitehead Institute for Biomedical Research, Nine Cambridge Center, and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, USA

† Department of Biology, University of California, B-022 Bonner Hall, San Diego, La Jolla, California 92093, USA

MAJOR histocompatibility complex (MHC) class I molecules are integral membrane proteins present on virtually all vertebrate cells and consist of a heterodimer between the highly polymorphic α -chain and the β_2 -microglobulin ($\beta_2\text{-m}$) protein of relative molecular mass 12,000 (ref. 1). These cell-surface molecules play a pivotal part in the recognition of antigens, the cytotoxic response of T cells, and the induction of self tolerance^{1,2}. It is possible, however, that the function of MHC class I molecules is not restricted to the immune system, but extends to a wide variety of biological reactions including cell–cell interactions. For example, MHC class I molecules seem to be associated with various cell-surface proteins, including the receptors for insulin, epidermal growth factor, luteinizing hormone and the β -adrenergic receptor^{3–6}. In mice, class I molecules are secreted in the urine and act as highly specific olfactory cues which influence mating preference^{7,8}. The $\beta_2\text{-m}$ protein has also been identified as the smaller

TABLE 1 Frequency of gene targeting in D3 cells

Expt	PCR ⁺ pools*	PCR ⁺ pools confirmed by cell cloning*		Targeted clones/total clones	Estimated targeting frequency‡
		G418 ^r colonies†			
1	7/32	3/3	15	5/31	1 in 30
2	19/28	4/4	36	5/40	1 in 16

ES cells were electroporated with the targeting vector and selected for G418^r as described in Fig. 2.

* Expressed per total number tested.

† Average number per 48-well plate.

‡ The frequency of targeting was calculated using the Poisson distribution formula, in which the fraction of plates with no recombinants, $P(0)$ is given by e^{-m} , where m is the average number of recombinants per pool ($m = -\ln P(0)$). The targeting efficiency is the ratio of m to the average number of independent G418^r clones per pool, which is half the observed number based on the assumption that each pool has on average two sibling clones. PCR⁺, scored positive by PCR.

component of the Fc receptor in neonatal intestinal cells⁹, and it has been suggested that the protein induces collagenase in fibroblasts¹⁰. Cells lacking $\beta 2\text{-m}$ are deficient in the expression of MHC class I molecules, indicating that the association with $\beta 2\text{-m}$ is crucial for the transport of MHC class I molecules to the cell surface¹. The most direct means of unravelling the many biological functions of $\beta 2\text{-m}$ is to create a mutant mouse with a defective $\beta 2\text{-m}$ gene. We have now used the technique of homologous recombination to disrupt the $\beta 2\text{-m}$ gene. We report here that introduction of a targeting vector into embryonic stem cells resulted in $\beta 2\text{-m}$ gene disruption with high frequency. Chimaeric mice derived from blastocysts injected with mutant embryonic stem cell clones transmit the mutant allele to their offspring.

Pluripotent embryonic stem cells can be genetically manipulated in culture, and when introduced into a host blastocyst, contribute to all tissues of the chimaeric mouse^{11,12}. The contribution of an embryonic stem (ES) cell carrying a defined mutation to the germ line of the chimaera thus results in a new mouse strain carrying the mutation. Gene targeting, mediated by recombination between a vector and homologous chromosomal sequences, has been used in ES cells to disrupt several genes, including the hypoxanthine-guanine phosphoribosyl transferase gene (*HPRT*)¹³⁻¹⁵, the *int-2* proto-oncogene¹⁴, the mouse *engrailed*-like gene *En-2* (ref. 16) and the *adipon*, the *aP2* and the *fos* gene²⁵. In these cases, the targeting vector contained the bacterial neomycin-resistance gene (*neo*^r) as a selectable marker gene facilitating the isolation of targeted clones. More recently, the herpes simplex virus (HSV) thymidine kinase gene (*tk*) has

been included in the vector to permit negative selection against nonhomologous integration events¹⁴. An alternative strategy using microinjection of a mutated gene fragment has been used to disrupt the *Hox1.1* gene in ES cells¹⁷. Recently, germ-line transmission of an *HPRT* mutant gene corrected by gene targeting has been reported¹⁸, confirming that gene disruption in ES cells can represent a general approach to generate defined mutations in mice.

To mutate the $\beta 2\text{-m}$ gene, a replacement-type vector was designed which combined several previously used characteristics^{13,14} facilitating the detection of targeting events. The vector contained 10 kilobases (kb) of homology to the $\beta 2\text{-m}$ gene, a *neo*^r cassette in the second exon lacking a polyadenylation site and a *tk* gene at the 5' end (Fig. 1a). Initial Southern-blot analyses revealed, however, that 50% of randomly picked G418-resistant (*Neo*^r) ES cell clones had lost the *tk* gene. Therefore, no counter selection against non-homologous integration events was applied in our experiments.

The following screening procedure was devised to allow rapid identification of targeted clones (Fig. 2). D3 ES cells¹² were electroporated on day 0, divided into ~30 independent pools and grown for 3.5 days in normal medium, by which time a cell containing a targeting event had divided into ~8 daughter cells. At that time, the pools of cells were split in half and subjected to drug selection. One-half of each pool (each containing 15-40 G418^r colonies after plating) was grown for 10 days as a bulk culture and used for polymerase-chain-reaction (PCR) analysis to identify the presence of a targeted cell. The other half of each

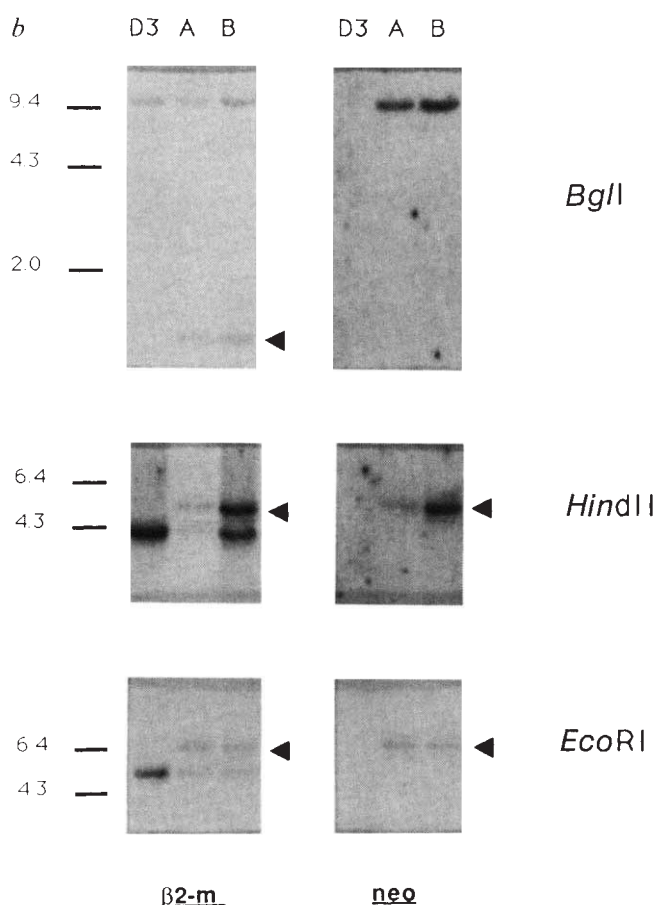
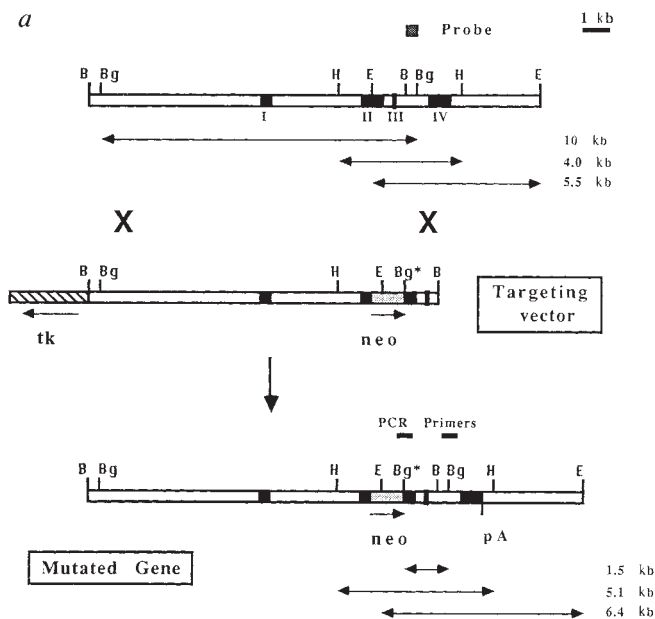


FIG. 1 a, Schematic diagrams showing the structure of the $\beta 2\text{-m}$ locus of parental D3 cells (top), the targeting vector (middle) and the predicted structure of a targeted $\beta 2\text{-m}$ locus (bottom). Not shown in the targeting vector are the 3.0-kb plasmid sequences upstream from the *tk* sequences. The *neo* cassette that we used contains the 1.1-kb blunt-ended *Xho*I-*Sall* fragment of plasmid pMCneo (ref. 13) inserted in the blunt-ended *Eco*RI site of exon II of the 10-kb *Bam*HI fragment of the $\beta 2\text{-m}$ gene. The *neo* gene in this construct is driven by the *tk* promoter with an upstream tandem repeat of the polyoma mutant enhancer region¹³. Small bars indicate the position of the primers used for PCR. A box indicates the position of the 500-base pair (bp) *Bam*HI-*Bgl*II ($\beta 2\text{-m}$) fragment used as hybridization probes in the PCR and/or Southern Blot analyses. The 900-bp *Eco*RI-*Sall* fragment used as a *neo* probe was derived from plasmid pMCneo (ref. 13). Roman numerals denote the exons of the $\beta 2\text{-m}$ gene. Arrows indicate the sizes of *Bgl*II, *Hind*III and *Eco*RI fragments hybridizing with the $\beta 2\text{-m}$ probe in DNA of parental D3 cells (top) and the predicted sizes in a targeted $\beta 2\text{-m}$ gene (bottom). Bg* is a *Bgl*II polymorphism present in exon II of the targeting vector (C57BL/6 DNA-derived, $\beta 2\text{-m}^b$ allele), but absent in the ES cell DNA (129J mouse-derived, $\beta 2\text{-m}^a$ allele). B, *Bam*HI site; Bg, *Bgl*II site; E, *Eco*RI

site; H, *Hind*III site. b, Southern-blot analyses of the structure of the $\beta 2\text{-m}$ locus in the DNA of parental D3 cells and two targeted clones A and B. Hybridization probes as in a. Triangles indicate diagnostic restriction enzyme fragments obtained after *Bgl*II, *Hind*III and *Eco*RI digestion. Bars indicate migration of *Hind*III-digested λ DNA markers. DNA (10 μ g) was digested with the indicated enzymes, run on a 0.7% agarose gel, and blot-hybridized by standard procedures.

TABLE 2 Germ-line transmission of ES cell genome

Animals from injected blastocysts		Progeny		Chimaeric mice		
Injected ES cells	No. blastocysts injected*	Born	Survived	Total tested†	Female	Male
Parental D3	18	4	4	4	2	2
Clone A	49	7	3	2	0	2
Clone B	40	17	8	8	0	8

Offspring of male chimaeras		Extent of coat-colour chimaerism†		No. of germ-line offspring with ES genome/per total‡ (% transmission)	
Chimaera derived from:	Mouse no.		No. of litters		
Parental D3	1	90%	5	21/33 (64%)	
	2	70%	3	17/28 (61%)	
Clone A	A1	5%	7	0/40 (0%)	
Clone B	B1	80%	5	1/53 (2%)	
	B2	95%	3	16/16 (100%)	
	B3	90%	1	7/7 (100%)	
	B4	50%	Sterile		
	B5	95%	5	22/22 (100%)	
	B6	90%	4	5/23 (22%)	
	B7	50%	3	4/19 (21%)	

* ES cells (10–20) were injected into C57BL/6J blastocysts and transferred to uteri of pseudopregnant recipients.

† Assessed by the contribution of the agouti coat colour contributed by D3 cells (129J, agouti) on the black background of C57BL/6J mice.

‡ Assessed by the presence of the dominant agouti coat colour in offspring after breeding with C57BL/6J females.

GENE TARGETING PROTOCOL

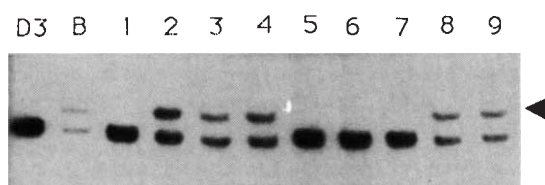
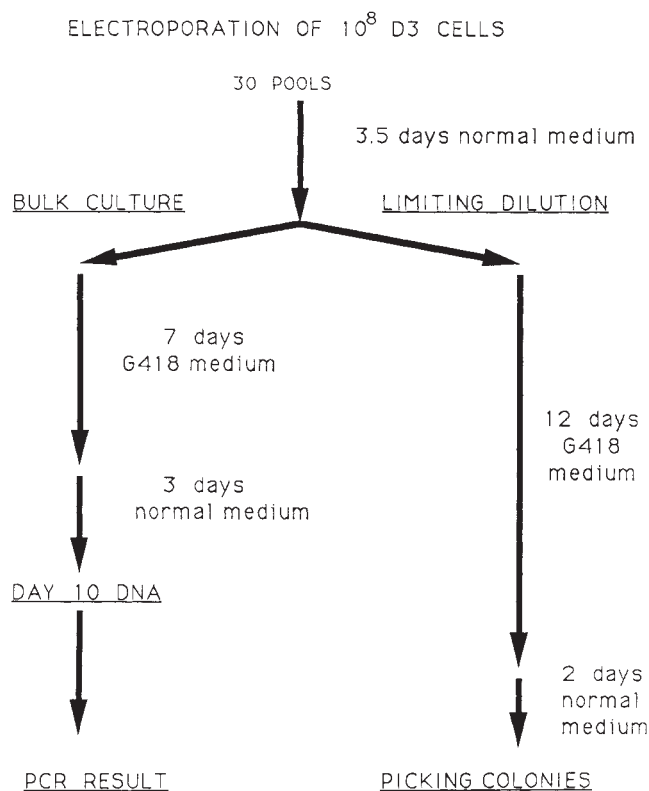


FIG. 2 Electroporation of ES cells. D3 ES cells¹² were routinely grown on γ -irradiated (2,000 rad) primary embryonic fibroblasts (EF) in DMEM medium supplemented with 15% FCS, $1 \times$ non-essential amino acids (Gibco) and 10^{-4} M β -mercaptoethanol. Electroporation of 1×10^8 D3 cells using a BTX 300 transfecter (50 μ F, 275 V) with 25 μ g ml⁻¹ linearized vector DNA was performed as previously described¹³. The electroporation conditions result in an 80% plating efficiency in comparison with non-electroporated D3 cells. Electroporated cells were divided into 30 independent pools and seeded into individual 25-cm² tissue culture flasks on a feeder layer of EF. Selection procedure. After three one-half-day's growth in normal medium, cells of individual pools were trypsinized and divided into two halves. One half was seeded into a 48-well plate on irradiated EF derived from transgenic mice expressing a *tk-neo* construct¹² and selected for 12 days in medium containing 150 μ g ml⁻¹ (active substance) G418. From day 6–14, the medium was supplemented with 15% D3 cell-conditioned medium (0.22- μ m filtered) and 1,000 U ml⁻¹ recombinant mouse leukaemia inhibitory factor (LIF)²⁴. The other half was seeded as a bulk culture in a 25-cm² flask on the irradiated *neo*-expressing EF. The bulk cultures were selected for 7 days in medium containing 150 μ g ml⁻¹ G418 and subsequently for 3 days in normal medium. Thereafter, DNA derived from the bulk cultures were tested by PCR for the presence of gene-targeted clones. Subsequently, the sibling cultures in the 48-well dishes were used to isolate clones from those pools scored positive by PCR. Clones were expanded and screened by PCR. Thereafter, DNA of PCR⁺ clones was characterized by Southern-blot analyses as outlined in Fig. 1. For PCR, DNA from bulk cultures or individual clones was isolated, treated with proteinase K, deproteinized with phenol-chloroform, precipitated and washed with ethanol and dissolved in 10 mM Tris buffer pH 8.0, 1 mM EDTA by standard procedures. DNA (250 ng) was tested by PCR under conditions recommended by the supplier (Perkin Elmer Cetus) with the addition of 2 mM MgCl₂ in 50 μ l containing 200 pmol of each primer (*neo*, ATATTGCTGAAGAGCTTGGCGGCGAATGGG; β 2-*m*, AAGGGAGGGAGAGAAGGAGAAGGTTAGCC). PCR was run for 40 cycles using a thermal cycler (Perkin Elmer Cetus). Denaturation was performed for 1.5 min at 94 °C, annealing for 2 min at 63 °C and extension for 4 min at 72 °C. A 25- μ l reaction sample was run on an 0.8% agarose gel and blot-hybridized by standard procedures.

FIG. 3 Germ-line transmission of mutant *b2-m* gene. Southern blot of DNA from offspring of chimaeric males B2 and B3 (Table 2). The DNA was digested with *Eco*RI and probed with a β 2-*m* probe as described in Fig. 1. The arrowhead marks the fragment from the mutant β 2-*m* gene (6.4 kb, see Fig. 1). D3, parental ES line; clone B, mutant clone injected into blastocysts; numbers refer to individual offspring.

pool was seeded into 48-well plates and propagated for 12 days in the presence of G418. As soon as pools containing targeted cells were identified by PCR analysis of the bulk cultures, clones were isolated from the corresponding sibling 48-well cultures, expanded and screened by PCR. The successful identification of targeted clones by this screening protocol depends critically on the plating efficiency of the trypsinized cells at day 3.5 after the initial electroporation. The use of a low trypsin concentration (0.025%)¹⁹ and the seeding of the cells at high density ($2-4 \times 10^5$ cells cm^{-2}) resulted in a plating efficiency of $>50\%$, thereby permitting partition of daughters of a single initially targeted cell into both sibling pools.

The results of two independent experiments are summarized in Table 1. In the first experiment, 7/32 cell pools, and in the second experiment, 19/28 cell pools, were identified to contain targeted cells by PCR. A large fraction of clones present in each of a total of seven positive pools were isolated, and targeted clones were identified by the PCR reaction in each pool. To confirm that the $\beta 2-m$ gene in the targeted clones was disrupted, DNA from seven independent clones was characterized by Southern-blot analysis using probes specific for the $\beta 2-m$ and *neo*^r genes (Fig. 1b). DNA of six clones showed the additional diagnostic *Bgl*II, *Eco*RI and *Hind*III fragments predicted from disruption of the $\beta 2-m$ gene by the targeting vector (Fig. 1a). DNA of one clone contained the diagnostic *Eco*RI and *Hind*III fragments, but lacked the *Bgl*II site in exon II (data not shown), most probably caused by a small mutation or deletion which can occur during a homologous recombination event¹⁵.

The targeting frequency in our experiments was ~ 1 in 2.5×10^6 electroporated input cells or 1/25 of the G418^r clones. This is one to three orders of magnitude higher than the targeting frequencies reported for other genes^{13-16,25}. There are at least two factors that could contribute to the high efficiency of $\beta 2-m$ -gene targeting. The gene could contain a hotspot of recombination, as suggested by the clustering of rearrangements in the first intron in the DNA of independent $\beta 2-m$ gene variants in lymphoma cell lines²⁰. In addition, the lack of a polyadenylation site in the vector reduced the number of drug-resistant clones from non-homologous integration events by about fourfold (data not shown).

Northern-blot analysis indicated the presence of little, if any, $\beta 2-m$ -specific messenger RNA in undifferentiated D3 cells (data not shown). Therefore, high-frequency homologous recombination in ES cells is clearly not dependent on a high steady-state mRNA level as exemplified in the study described here for the $\beta 2-m$ gene, and observed previously for other genes^{16,25}. In our view, the chromatin structure could be more important. It has been shown, for example, that the $\beta 2-m$ gene in F9 cells, which express only trace amounts of $\beta 2-m$ mRNA, is much more susceptible to digestion by DNAase I than it is in differentiated F9 cells, which express moderate levels of $\beta 2-m$ mRNA²¹. This suggests that the gene, even if it is expressed a little or not at all is in an 'opened' chromatin conformation.

To derive animals carrying a disrupted $\beta 2-m$ gene, two different targeted clones were injected into C57BL/6J host blastocysts. Whereas clone A resulted in a few animals with a low degree of chimaerism, extensive and healthy chimaeras were derived from clone B (Table 2). Chimaeric males were bred with C57BL/6J females. So far, three males have transmitted the ES cell genome to all of their offspring and three other males have transmitted it to part of their offspring (Table 2). A 100% transmission is expected for 'pseudomales', which are derived from a female host blastocyst injected with the male ES cell line²². In such animals, the XX germ cells of the female host are not capable of forming functional spermatozoa. So far, Southern-blot analysis was performed with tail DNA of 17 offspring from pseudomales B2 and B3 (Table 2), and revealed the presence of the additional *Eco*RI fragment predicted by inheritance of the disrupted $\beta 2-m$ gene in eight animals (Fig. 3). Our results, therefore, indicate that the disrupted $\beta 2-m$ gene

is genetically transmitted according to mendelian expectations.

The animals heterozygous for the $\beta 2-m$ -gene disruption have so far not shown any detectable phenotype and will be intercrossed to obtain animals homozygous for the mutation. We expect that homozygous animals will lack MHC class I molecules on their cell surface and, as a consequence, be severely immunocompromised. Considering the various other functions the $\beta 2-m$ protein could have, failure to express the gene may lead to disturbance of embryonic development. The $\beta 2-m$ gene is expressed as soon as the two-cell stage²³, and thus is one of the earliest genes to be activated during mammalian development that is known. The phenotype of homozygous mutant mice should therefore help in understanding the various functions that the $\beta 2-m$ protein, as well as its ligands, has in the life cycle of mammals. □

Received 18 September; accepted 20 October 1989

1. Klein, J., Juretic, A., Baxevas, C. N. & Nagy, Z. A. *Nature* **291**, 455-460 (1981).
2. Zinkernagel, R. M. & Doherty, P. C. *Adv. Immun.* **27**, 51-117 (1979).
3. Fehlmann, M. et al. *Proc. natn. Acad. Sci. U.S.A.* **82**, 8634-8637 (1985).
4. Due, C., Simonsen, M. & Olsson, L. *Proc. natn. Acad. Sci. U.S.A.* **83**, 6007-6011 (1986).
5. Schreiber, A. B., Schlessinger, J. & Eddins, M. J. *J. cell. Biol.* **98**, 725-731 (1984).
6. Solano, A. R. et al. *Proc. natn. Acad. Sci. U.S.A.* **85**, 5087-5091 (1988).
7. Singh, P. B., Brown, R. E. & Roser, B. *Nature* **327**, 161-164 (1987).
8. Yamazaki, K. et al. *Science* **240**, 1331-1332 (1988).
9. Simister, N. E. & Mostov, K. E. *Nature* **337**, 184-187 (1989).
10. Brinckerhoff, C. E., Mitchell, T. I., Karmilowicz, M. J., Kluge-Beckerman, B. & Benson, M. D. *Science* **243**, 655-657 (1989).
11. Robertson, E., Bradley, A., Kuehn, M. & Evans, M. *Nature* **323**, 445-448 (1986).
12. Gossler, A., Doetschman, T., Korn, R., Serfling, E. & Kemler, R. *Proc. natn. Acad. Sci. U.S.A.* **83**, 9065-9069 (1986).
13. Thomas, K. R. & Capecchi, M. R. *Cell* **51**, 503-512 (1987).
14. Mansour, S. L., Thomas, K. R. & Capecchi, M. R. *Nature* **336**, 348-352 (1988).
15. Doetschman, T., Maeda, N. & Smithies, O. *Proc. natn. Acad. Sci. U.S.A.* **85**, 8583-8587 (1988).
16. Joyner, A. L., Skarnes, W. C. & Rossant, J. *Nature* **338**, 153-156 (1989).
17. Zimmer, A. & Gruss, P. *Nature* **338**, 150-153 (1989).
18. Thompson, S., Clarke, A. R., Pow, A. M., Hooper, M. L. & Melton, D. W. *Cell* **56**, 313-321 (1989).
19. Rudnicki, M. A. & McBurney, M. W. in *Teratocarcinomas and Embryonic Stem Cells* (ed. Robertson, E. J.) 19-49 (IRL, Oxford, 1987).
20. Potter, T. A., Frankel, W., Zeff, R. A. & Rajan, T. V. *J. Immun.* **138**, 1270-1274 (1987).
21. Croce, C. M. et al. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5754-5758 (1981).
22. Bradley, A., Evans, M., Kaufman, M. & Robertson, E. *Nature* **309**, 255-256 (1984).
23. Sawicki, J. A., Magnuson, T. & Epstein, C. J. *Nature* **294**, 450-451 (1981).
24. William, R. L. et al. *Nature* **336**, 684-687 (1988).
25. Johnson, R. et al. *Science* **245**, 1234-1236 (1989).

ACKNOWLEDGEMENTS. We thank Dr Tom Gridley for his assistance in establishing the culture of ES cells and the blastocyst injections. We thank Drs Achim Gossler and Rolf Kemler for the *tk-neo* transgenic mice and D3 cells, Nicholas Gough for providing us with LIF and Philippe Kourilsky for his gift of the cB2-m2 cosmid. We thank Drs Michael Rudnicki, William Strauss, Kuo-Fen Lee and Neil Simister for discussions and critical reading of the manuscript. This work was supported by the Netherlands Organization for the Advancement of Pure Research (M.Z.), a Cancer Research Institute/Fan Fox & Leslie R. Samuels Foundation Fellowship (M.Z.) and the NIH (S.S. and R.J.).

Developmental regulation of stage-specific ribosome populations in *Plasmodium*

A. P. Waters, Chiang Syin* & T. F. McCutchan†

Malaria Section, Laboratory of Parasitic Diseases, National Institute of Allergy and Infectious Disease, Building 4, National Institutes of Health, Bethesda, Maryland 20892, USA

THE *Plasmodium* parasites are so far unique in biology in possessing developmentally regulated ribosomal rRNA gene units¹. Two different genes encode their small subunit rRNAs: one gene (A) yields transcripts predominant in the asexual blood-stage parasites, and the other (C) is mainly transcribed in the sporozoite forms that develop in the mosquito^{1,2}. Developmental control of events allowing a switch in the complement of ribosomes must coordinate the production of the new class with selective inactivation and removal of the old. We show here that in *P. falciparum* the switch from A to C gene expression involves the control of

* Present address: Division of Biochemistry and Biophysics, Center for Biologics Evaluation, and Research, Food and Drug Administration, Bethesda, Maryland 20892, USA.

† To whom correspondence should be addressed.

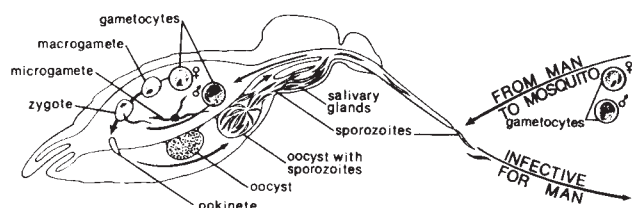


FIG. 1 Progression of sexual development of *Plasmodium*.

rRNA processing, allowing accumulation of precursor *C*-gene transcripts in gametocytes. These precursor molecules are processed to mature size in the zygote and the early ookinete, where both transcription and processing of the *C*-gene rRNA seem to be accelerated. As the *C*-gene precursor rRNA appears, a defined and limited pattern of breakdown of the dominant *A*-gene rRNA occurs, in which conserved, functionally active sequences involved in the termination of translation and elongation are targeted. By the late oocyst stage¹, the *A*-gene transcripts are virtually replaced by mature *C*-gene transcripts.

A generalized scheme showing the malaria parasite's sexual development is shown in Fig. 1. The parasites are transmitted from the vertebrate host to the mosquito in the form of gametocytes which are contained in the blood meal; the gametocytes mature to extracellular gametes in the midgut of the mosquito. The gametes then fertilize and the resulting zygotes mature to a motile stage, the ookinete, which transverse the mosquito midgut wall and continues to develop. Here we follow the course of transition between the ribosome types during this part of the malaria life cycle, given that at different points in the life cycle either the *A* or the *C* gene is dominantly expressed. We investigated the nature of the switch during gametocytogenesis that facilitates the change between *A* gene expression in asexual parasites and *C* gene expression in late oocysts and sporozoites¹.

We analysed equal amounts of total RNA from gametocytes, zygotes and ookinetes by northern blotting. In Fig. 2, identical blots were hybridized with a probe specific either for *A*-gene transcripts (TM40; Fig. 2, right panel) or for those from the *C* gene (TM39; Fig. 2, left panel). The mature *A*-gene transcript is dominant in all three stages. This gene seems to be actively transcribed in gametocytes and zygotes, where small amounts of its precursors are seen, whereas the number of *A*-gene precursor transcripts observed in ookinetes is reduced. The *C* gene-specific transcripts are almost entirely in a precursor form which is ~1,800 nucleotides larger than the mature small subunit RNA in gametocytes. This precursor is also extremely stable, accumu-

lating over the duration of gametocytogenesis. Processing of this precursor is initiated in zygotes, and in 24-hour ookinetes most of the *C*-specific RNA is observed as the mature form. Also, the relative amount of the *C*-gene transcript increases. Considered with the observed decrease in the amount of mature *A*-gene transcription, the data indicate that there has been a change in the relative amounts of the two rRNA transcripts.

The established pattern of small-subunit rRNA gene expression¹ suggests that, at this point in the life cycle, the parasite must not only increase the amount of *C*-gene transcript, but also reduce the level of mature *A*-gene transcripts, which in the gametocyte probably represent 90–95% of total RNA. We have previously shown that by the eighth day of development of the oocyst, *A*-gene transcripts have virtually disappeared¹. The data shown in Fig. 3a indicate that the process of turnover specific to the *A*-gene transcript must involve at least two site-specific cleavage events, rather than a generalized breakdown by nuclease action. We used oligonucleotide mapping on northern blots to demonstrate that the first site of cleavage lies in a 10-nucleotide region between positions 1,664 and 1,673 (Fig. 3b). Nearly 30% of the signal hybridizing to gametocyte RNA appears in a processed fragment that is ~300 nucleotides long; 70% remains in the apparently unprocessed mature RNA. The relative proportion of this 300-nucleotide fragment in zygote RNA has increased to almost 40% of the total. In the 24-hour ookinete RNA, the 300-nucleotide fragment is secondarily cleaved to yield a new fragment ~150 nucleotides in length. Clearly the sequence of events is ordered. Little, if any, processing of this type is seen in RNA from asexual parasites (data not shown). The site of the initial cleavage of the *A*-gene RNA is interesting in that it lies within core sequence, which implies that cleavage affects a functional region of the sub-unit and presumably inactivates the ribosome. The equivalent region in the *Escherichia coli* small-subunit rRNA sequence has been implicated in UGA termination^{3,4}. Although the sequence of this region differs in prokaryotes and eukaryotes, it is conserved within the two kingdoms⁵. In this respect, the important cytosine residue (at *E. coli* position 1,054; ref. 3) is universally conserved⁵. The second coordinate cut occurs very close to the decoding site within the subunit. The specific location of these cuts is of interest, given that many areas of the ribosome will be accessible, yet both cuts occur at precise locations possibly within functional areas. The site of the initial cut might be more accessible in dissociated subunits than in actively translating ribosomes, so this could provide a means of discrimination for the nuclease and control over the limited nature of the *A*-gene ribosome degradation.

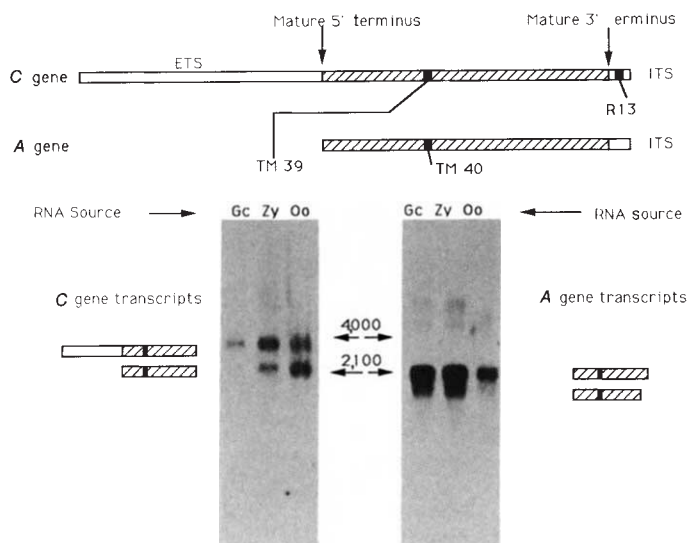


FIG. 2 *Plasmodium* produces a stable precursor of the *C*-gene small-subunit rRNA during sexual development. Gametocytes (Gc), Zygotes (Zy) and ookinetes (Oo) of *P. falciparum* were produced by *in vitro* culture⁷. The term ookinete is used to describe the culture of zygotes after 24 h. Although expression of zygote surface antigens has ceased morphologically, the parasite form is not the classical ookinete. We designate this cultured material ookinete to indicate that some clear morphological development of the material beyond zygote has occurred. RNA was isolated by standard methods⁸ and analysed on a 1.0% agarose gel containing formaldehyde. After staining and photography, the RNA was transferred for northern analysis to a Hybond-N membrane (Amersham). The probes were complementary to the designated regions: TM39 is complementary to nucleotides 669–685 in the *C*-gene small-subunit rRNA sequence, and TM40 is complementary to nucleotides 672–687 in the *A*-gene small-subunit rRNA sequence. These sequences have been reported elsewhere². Oligo R13 is complementary to the *C* gene internal transcribed spacer (A.P.W. unpublished data) at positions 17 to 36 (numbering is from the mature 3' terminus) and was used to demonstrate that the precursor has the same terminus as the mature small-subunit rRNA. Its sequence is 5'-TACATTTCTTTCTTTCTTA-3'. All hybridizations were under standard conditions⁹. Washing included a final stringency wash in 2×SSC, 1% SDS at 5°C below the calculated T_m . The migration positions of the large and small rRNAs (sizes in nucleotides) are indicated. ETS, External transcribed spacer. ITS, Internal transcribed spacer.

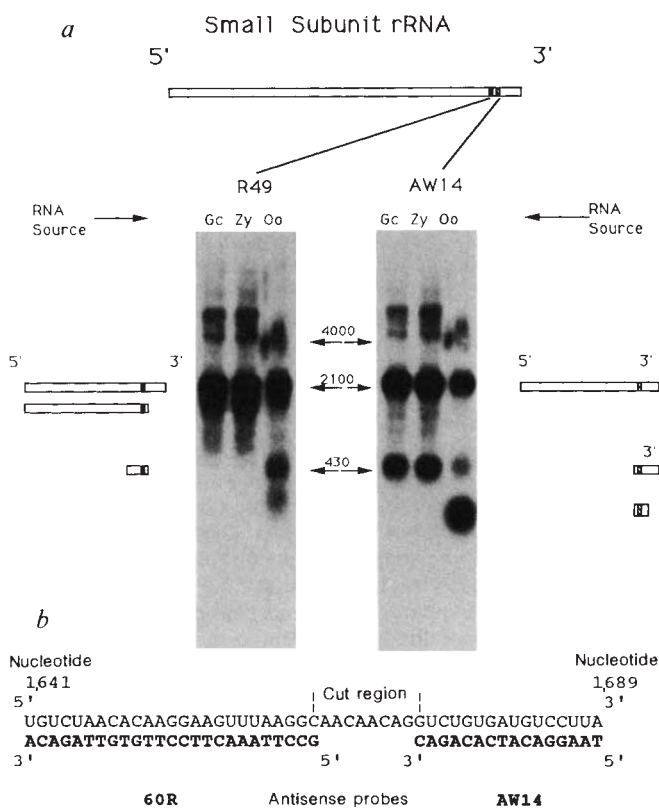


FIG. 3 *a*, The A-gene small-subunit rRNA transcript is selectively cleaved during sexual development. The probes were complementary to the regions indicated: AW14 has the sequence 5'-TAAGGACATCAGAC-3' and is complementary to nucleotides 1,674–1,689; probe 49R has been described⁹ and is complementary to nucleotides 1,606–1,634; probe 60R has the sequence 5'-GCCTTAAACTTCCTGTGTAGACA-3' and was used to map the precise location of the initial cut. Methods are described in the legend to Fig. 2. *b*, Site of the initial cut in the *P. falciparum* A gene. The sequence of the A-gene small-subunit rRNA region between nucleotides 1,641–1,689 is given at the top. The two antisense oligonucleotide probes, AW14 and 60R, are shown in bold at the appropriate locations. The two probes hybridize at stringency to the two different processed RNA fragments in gametes and gametocytes. The region of the first cut site is thus located to a region of no more than 10 nucleotides between the two probes.

The structure of the stable precursor in gametocytes from the ribosomal C gene is also significant to the understanding of the transition process. Oligonucleotide mapping has demonstrated that the entire 1,800 extra nucleotides in the precursor lies 5' to the mature transcript (data not shown). Such a precursor structure is like that of small-subunit rRNA precursor molecules found in yeast mutants that lack the gene for a ubiquitinated ribosomal protein⁶. It is possible that a stage- (and binding sequence-) specific ubiquitinated ribosomal protein is produced to aid the maturation of each type of rRNA transcript; this is under investigation at present.

Although much is known about the generation of mature rRNA transcripts from the single multigenic precursor molecule, our results demonstrate that it can be a regulated, developmentally controlled process. This control is minimally exerted at three levels, transcription, RNA processing and RNA degradation, facilitating the novel phenomenon of ribosome biogenesis, switching and turnover in *Plasmodium*. Specific factors involved in these events during the hepatic stage of the *Plasmodium* life cycle could prove to be targets for novel drugs directed against malaria parasites and other micro-organisms. These events can be easily assayed *in vitro* to test such drugs. It is possible that other organisms with a defined dimorphic lifestyle could also display this type of control. □

Received 12 July; accepted 9 October 1989.

- Gunderson, J. H. *et al.* *Science* **238**, 933–937 (1987).
- McCutchan, T. F. *et al.* *Molec. Biochem. Parasitol.* **28**, 63–68 (1988).
- Murgola, E. J., Hijazi, K. A., Goring, H. U. & Dahlberg, A. E. *Proc. natn. Acad. Sci. U.S.A.* **85**, 4162–4165 (1988).
- Dahlberg, A. E. *Cell* **57**, 525–529 (1989).
- Dams, E. *et al.* *Nucleic Acids Res.* **16**, r87–r173 (1988).
- Finley, D., Bartel, B. & Varshavsky, A. *Nature* **338**, 394–401 (1989).
- Carter, R., Graves, P. M., Quakyl, I. A. & Good, M. F. *J. Exp. Med.* **169**, 135–147 (1989).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* 194–195 (Cold Spring Harbour Laboratory, New York, 1982).
- Waters, A. P. & McCutchan, T. F. *Lancet* **i**, 1343–1346 (1989).

ACKNOWLEDGEMENT. We thank Dr David Kaslow for discussion and for critical reading of the manuscript.

Molecular cloning and expression of human hepatocyte growth factor

Toshikazu Nakamura*, Tsutomu Nishizawa†, Mitsuo Hagiya†, Tatsuya Seki†, Manabu Shimonishi†, Atsushi Sugimura†, Kosuke Tashiro* & Shin Shimizu†

* Department of Biology, Faculty of Science, Kyushu University, Fukuoka 812, Japan

† Pharmaceuticals Research Center, Toyobo Co. Ltd, Ohtsu 520-02, Japan

HEPATOCTE growth factor (HGF) is the most potent mitogen for mature parenchymal hepatocytes in primary culture, and seems to be a hepatotrophic factor that acts as a trigger for liver regeneration after partial hepatectomy and liver injury. The partial purification and characterization of HGF have been reported^{1–4}. We have demonstrated that pure HGF from rat platelets is a new growth factor^{5,6} effective at concentrations as low as 1 ng ml⁻¹.

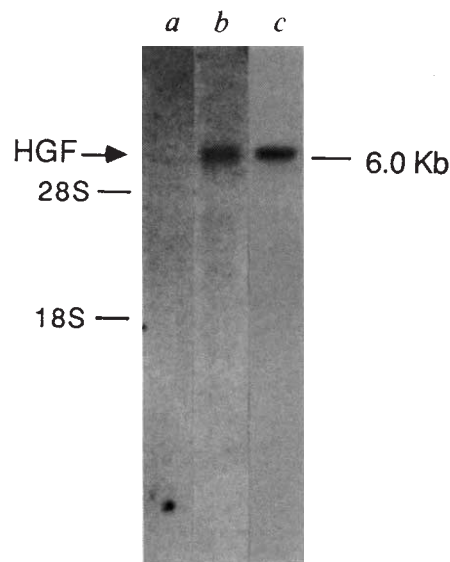


FIG. 1 Northern hybridization of mRNAs from rat and human livers. Poly(A)⁺ RNAs were isolated from livers of normal adult rat (lane *a*) and CCl₄-treated rat (lane *b*). Lane *c* is human liver poly(A)⁺ RNA (Clontech). The positions of the 28S and 18S ribosomal RNA markers are shown on the left. Blots were hybridized with rat HGF cDNA (clone RBC1: 1.4 kb insert).

METHODS. Poly(A)⁺ RNAs (1 µg per lane) were electrophoresed into a 1% agarose gel containing 0.66 M formaldehyde¹⁷ and blotted onto a nylon membrane filter (Genescreen plus; NEN). The filter was hybridized in 5× SSPE (800 mM NaCl, 50 mM sodium phosphate, 5 mM EDTA, pH 7.8), 5× Denhardt's solution, 10% dextran sulphate, 40% formamide, 0.1% SDS, 1 mM sodium pyrophosphate, 0.1 mg ml⁻¹ sonicated *E. coli* DNA at 40 °C overnight with a ³²P-labelled probe prepared from the 1.4 kb *Eco*RI fragment of the cDNA clone RBC1 (Amersham multiprime DNA labelling system). The filter was washed with 1× SSC containing 0.1% SDS at 60 °C, dried and then exposed on X-ray film for 2 days at -80 °C using an intensifying screen.

corresponding to a recovery of 23%. The purified HGF was reduced, carboxymethylated and fractionated by reverse-phase HPLC on a C_4 -column to isolate the α - and β -chains. We determined the N-terminal sequence (one-letter code) of the α - and β -chains as PLVKIKTKKVNSADECANR and VVNGIPTQTTVGWMVSLKYRNKHICGG, respectively.

We employed the following strategy to isolate a cDNA encoding complete human HGF. First, we cloned rat HGF cDNA using a hybridization probe designed to have a sequence corresponding to the N-terminus of the rat HGF β -chain, and then isolated human HGF cDNA using the rat HGF cDNA as probe. We investigated the tissue distribution of HGF in rats to establish a good source of HGF messenger RNA and found that the livers of rats treated with CCl_4 contain large amounts of HGF. To clone HGF cDNA, a cDNA library constructed in λ gt10 from poly(A)⁺ RNA extracted from the livers of CCl_4 -treated rats was screened with a single 44-mer oligodeoxynucleotide whose sequence corresponded to the first 15 amino acids of the β -chain. Screening 2×10^6 clones gave one positive clone with a 1.4-kilobase (kb) insert (clone RBC1). Nucleotide sequencing of this insert revealed an open reading frame predicting the translation of a 404-residue protein and a 3'-untranslated sequence (data not shown). We confirmed that the 27 N-terminal amino

acids of the β -chain were included in the 233 amino acids of the rat HGF β -chain encoded by the 1.4 kb insert.

Northern hybridization using the RBC1 clone showed that levels of rat HGF mRNA were negligible in normal liver but increased markedly in the injured livers of CCl_4 -treated rats, and that the mRNA was ~6 kb in size. We next investigated whether we could use human liver mRNA as a source of mRNA for human HGF cDNA cloning: HGF mRNA was clearly detectable in human liver and was the same size as the rat HGF mRNA, as shown in Fig. 1. A human liver cDNA library was therefore constructed in λ gt10, using oligo(dT) primer. The screening of 2×10^5 cDNA clones with 32 P-labelled RBC1 gave two positive clones (HBC1 and HBC25) and two additional cDNA clones (HAC69 and HAC102) were obtained from the unique 18-mer primer-extended cDNA library. Four clones were characterized in detail and are illustrated in Fig. 2a; together they span 6 kb which is sufficient to encode the entire mRNA that we detected by northern blotting. The sequence comprises a 5' non-coding region of 134 nucleotides, a single open reading frame of 2,184 nucleotides and 3,580 nucleotides of the 3' non-coding region, which contains an A⁺U-rich sequence (Fig. 2b).

A polypeptide of 728 amino acids is encoded by human HGF cDNA in a single open reading frame which includes both the α - and β -chains. The calculated relative molecular masses (M_r) of the human HGF pre-precursor and the mature form are 83,126 and 76,879 respectively. The N-terminal amino acid of the α -chain is preceded by 54 amino acids, starting with a methionine. Of these 54 amino acids, most of the first 29 are hydrophobic, which is typical of a signal sequence⁷. The next 25 (amino acids 30–54) preceding the N-terminus of the α -chain constitute the pro-sequence. The α -chain runs from amino acids 55 to 494, and the β -chain from 495 to 728. From the known substrate specificity of proteases, this Gln-Leu-Arg-Val sequence should be accessible to a trypsin-like protease such as plasmin or plasma kallikrein⁸. After biosynthesis, the chains are separated by a trypsin-like protease at an Arg-Val site (Fig. 4a). The predicted M_r of the α -chain, which consists of 440 amino acids, is 50,808, and for the β -chain, which consists of 234 amino acids, it is 26,089. The HGF contains four putative asparagine-linked glycosylation sites, which are located at positions 294 and 402 of the α -chain and at positions 566 and 653 of the β -chain. The difference between the M_r s of the HGF subunits as determined by SDS-PAGE and those predicted from the cDNA sequence may be due to carbohydrate residues, which have been demonstrated by binding to concanavalin A-Sepharose columns (data not shown).

We demonstrated that biologically active HGF could be expressed from its cDNA in transient expression experiments in COS-1 cells. As shown in Fig. 3, the conditioned medium from the cells transfected with the plasmid pEV(hHGF) revealed the dose-dependent stimulation of DNA synthesis of adult rat hepatocytes in primary culture. This activity was abolished by heating at 100 °C for two minutes and we found that it was additive with insulin and EGF, which suggested to us that it was attributable to HGF.

We found considerable homology (38%) between the overall amino-acid sequence of the pro-HGF precursor and plasminogen⁹, but neither plasminogen nor plasmin has HGF activity (data not shown). Plasminogen has 791 amino acids and contains five kringle structures; the α -chain of HGF would contain four kringle structures. Prothrombin, tissue plasminogen activator, urokinase and coagulation factor XII also contain a kringle structure¹⁰. Although the kringle domain is thought to play a part in protein-protein interaction, its physiological function is unknown. On the other hand, the mature β -chain has extensive homology (37%) with a serine protease domain of plasminogen, but the histidine and serine residues of the protease active site are replaced in the HGF β -chain by glutamine and tyrosine, respectively, so HGF should have no proteolytic

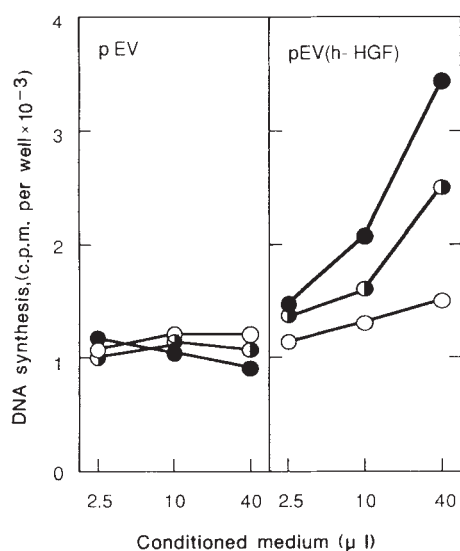


FIG. 3 Expression of the cloned cDNA encoding the human HGF in COS-1 monkey kidney cells. HGF activity in the conditioned medium from cultures of COS-1 cells transfected with HGF cDNA (right) and the corresponding vector without insert (left). ○, 1-day culture; ◐, 2-day culture; ●, 3-day culture.

METHODS. The vector pEV contains the SV40 early promoter, polyadenylation sequence and the origin of replication. The human HGF cDNA (*Bam*HI–*Pst*I 2.95 kb fragment) was inserted into pEV vector downstream of the SV40 early promoter. The constructed plasmid DNAs (pEV) were introduced to COS-1 cells by the DEAE-Dextran method²³. COS-1 cells were seeded at a density of 2×10^5 cells into 60 mm tissue culture plates two days before transfection. After washing the plates with 20 mM Tris-HCl (pH 7.4) containing 150 mM NaCl (TBS), 10 μ g plasmid DNA in TBS (600 μ l) containing 500 μ g ml⁻¹ DEAE-dextran was used for transfection for 30 min at room temperature. The transfected cells were incubated with Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal calf serum (FCS) and 100 μ M chloroquine for 3 h at 37 °C, followed by treatment with DMEM containing 10% FCS and 10% dimethylsulphoxide for 2.5 min at room temperature. This medium was replaced with DMEM containing 10% FCS and 50 μ g ml⁻¹ gentamicin and cultured for 3 days at 37 °C. The culture medium was collected at 24, 48 and 72 h after transfection, and HGF activity assayed as described^{24,25}. Aliquots of conditioned medium (2.5, 10 or 40 μ l) were added to the adult rat hepatocyte cultures (1.6 cm diameter well: 0.5 ml medium) 20 h after plating, and 18 h later, 125 I-labelled deoxyuridine (0.3 μ Ci ml⁻¹) was added. Six hours later, DNA synthesis was assayed by measuring the incorporation of 125 I-labelled deoxyuridine. A purified preparation of rat HGF from platelets (2 ng ml⁻¹) and saline used as controls gave 5,400 c.p.m. and 1,200 c.p.m. incorporation, respectively.

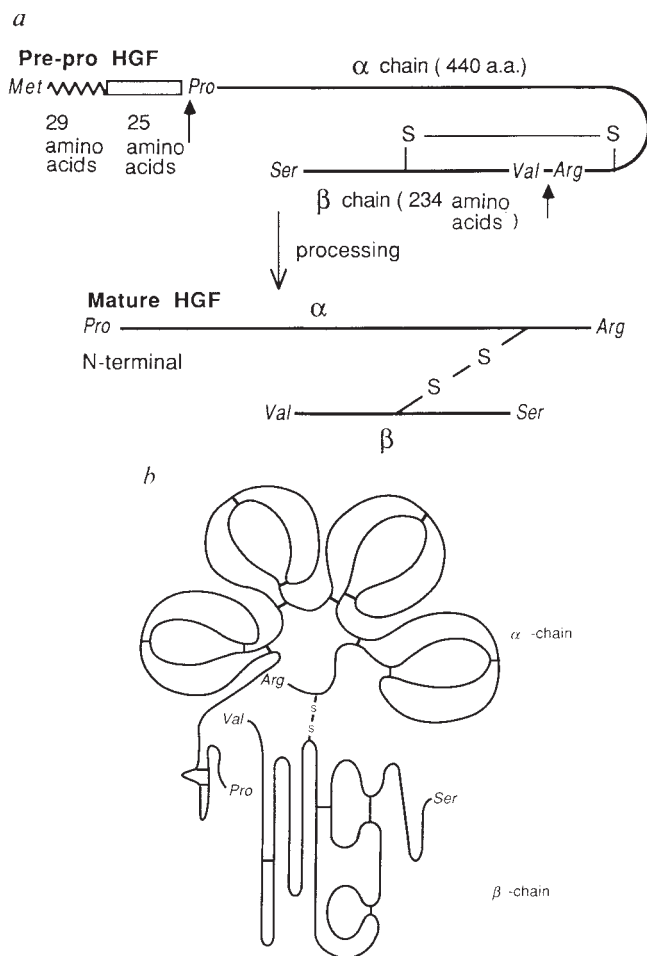


FIG. 4 Structural features of human HGF. *a*, Schematic representation of the structure and processing of pre-pro HGF. The arrows indicate the processing sites. The wavy line and unfilled line represent the signal peptide and pro-sequences, respectively. *b*, Predicted primary structure of human HGF. An interchain S-S bridge is formed between Cys 487 in the α -chain and Cys 604 in the β -chain.

activity. Other examples of amino acid replacements in the serine protease active site include haptoglobin¹¹ and protein Z¹², which have no proteolytic activity. In mature HGF, Cys 604 in the β -chain and Cys 487 in the α -chain seem to form an interchain bridge in a position comparable to that of the interchain bridge found in plasmin. The predicted structure of human HGF and an interchain S-S bridge is shown in Fig. 4b.

A marked increase in HGF-like activity has been reported in the sera of both partially hepatectomized^{1,2,13} and hepatotoxin-treated rats^{14,15} and of patients with fulminant hepatic failure¹⁶. We find a time-dependent increase in HGF activity in rat liver after administration of CCl₄, which reaches a maximum of about 20 times the initial level at 36 hours. The level of HGF mRNA is negligible in normal rat liver, but increases dramatically in injured liver 10 hours after the administration of CCl₄. We have demonstrated that rat liver-derived HGF has the same chemical and biological properties as rat platelet-derived HGF (manuscript in preparation). These findings indicate that the liver itself is producing HGF. HGF in plasma from patients with fulminant hepatic failure is probably derived from the injured liver as human platelets do not contain a significant amount of HGF, unlike rat platelets. Thus HGF seems to be an important growth factor in the stimulation of liver regeneration after partial hepatectomy and liver injury.

The availability of functionally active recombinant human HGF will enable us to study the HGF receptor on hepatocytes and the role of HGF during liver regeneration *in vivo*. HGF

may eventually be useful clinically for enhancing liver regeneration after surgery and in the diagnosis and treatment of hepatitis. □

Received 27 July; accepted 5 October 1989.

1. Nakamura, T., Nawa, K. & Ichihara, A. *Biochem. biophys. Res. Commun.* **122**, 1450-1459 (1984).
2. Thaler, F. J. & Michalopoulos, G. K. *Cancer Res.* **45**, 2545-2549 (1985).
3. Ghoda, E. *et al. J. clin. Invest.* **81**, 414-419 (1988).
4. Russell, W. E., McGowan, J. A. & Bucher, N. R. L. *J. Cell Physiol.* **119**, 193-197 (1984).
5. Nakamura, T., Teramoto, H. & Ichihara, A. *Proc. natn. Acad. Sci. U.S.A.* **83**, 6489-6493 (1986).
6. Nakamura, T., Nawa, K., Ichihara, A., Kaise, N. & Nishino, T. *FEBS Lett.* **224**, 311-316 (1987).
7. Von Heijne, G. *Eur. J. Biochem.* **133**, 17-21 (1983).
8. Kawabata, S. *et al. Eur. J. Biochem.* **172**, 17-25 (1988).
9. Sottrup-Jensen, L., Claeys, H., Zajdel, M., Petersen, T. E. & Magnusson, S. *Progress in Chemical Fibrinolysis and Thrombolysis* Vol. 3, 191-209 (Raven, New York, 1978).
10. Furie, B. & Furie, B. C. *Cell* **53**, 505-518 (1988).
11. Kurosky, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 3388-3392 (1980).
12. Højrup, P., Jensen, M. S. & Petersen, T. E. *FEBS Lett.* **184**, 333-338 (1985).
13. Morley, C. G. D. & Kingdom, H. S. *Biochim. biophys. Acta* **308**, 260-275 (1973).
14. Morley, C. G. D. & Boyer, J. L. *Biochim. biophys. Acta* **477**, 165-176 (1977).
15. Ando, T. *et al. Hepatology* (in the press).
16. Nakayama, H. *et al. Biomed. Res.* **6**, 231-237 (1985).
17. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* 202-203 (Cold Spring Harbor Laboratory, New York, 1982).
18. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* 125-126 (Cold Spring Harbor Laboratory, New York, 1982).
19. Gubler, U. & Hoffman, B. J. *Gene* **25**, 263-269 (1983).
20. Huynh, T. V., Young, R. A. & Davis, R. W. *DNA Cloning* Vol. 1, 49-78 (IRL, Oxford, 1985).
21. Benton, W. D. & Davis, R. W. *Science* **196**, 180-182 (1977).
22. Sanger, F., Miklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
23. Berger, S. L. & Kimmel, A. R. *Meth. Enzym.* **102**, 684-704 (Academic, London, 1987).
24. Nakamura, T., Tomita, Y. & Ichihara, A. *J. Biochem.* **94**, 1029-1035 (1983).
25. Nakamura, T., Arakaki, R. & Ichihara, A. *Exp. Cell Res.* **179**, 488-497 (1988).

ACKNOWLEDGEMENTS. This work was supported by the Ministry of Education and Culture of Japan and the Princess Takamatsu Cancer Research Fund.

Anti-HLA-A2 antibody-enhancement of peptide association with HLA-A2 as detected by cytotoxic T lymphocytes

Helen Bodmer, Graham Ogg, Frances Gotch & Andrew McMichael

Molecular Immunology Group, Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DU, UK

MOST cytotoxic T lymphocytes (CTL) not only recognize epitopes of viral or other foreign proteins in association with class I major histocompatibility complex (MHC) molecules, but also recognize target cells sensitized with short synthetic peptides representing the epitopes¹. There is increasing evidence that these synthetic peptides associate with the class I molecule both at the cell surface and intracellularly². We have now investigated the effect of a monoclonal antibody specific for HLA-A2 and HLA-B17 (B57/58) molecules (antibody MA2.1)³ on the sensitization of target cells with peptide for lysis by HLA-A2-restricted CTL. Previously, anti-HLA class I monoclonal antibodies have been shown to inhibit the recognition of target cells, infected with influenza A virus, by virus-specific CTL^{4,5}. We find, however, that target cells treated with MA2.1 antibody can be sensitized with peptide for CTL lysis much more rapidly than untreated cells, or at >100-fold lower peptide concentration than that required for sensitization of untreated cells. This implies that the antibody, which is believed to bind to one side of the peptide-binding groove, directly affects the binding of peptide to the HLA-A2 molecule at the cell surface.

Experiments using anti-HLA class I monoclonal antibodies to inhibit target-cell lysis by virus-specific CTL have demonstrated the importance of the class I molecule in MHC-restricted recognition of antigen^{4,5}. Most allele-specific antibodies are thought to bind to the helices on either side of the putative

peptide-binding groove of class I molecules⁶, thereby possibly effecting the binding of peptide to the class I molecules.

We investigated the effect of antibody MA2.1, a high-affinity antibody with specificity for HLA-A2 and HLA-B17 molecules^{3,7} on the association of peptide with HLA-A2-expressing target cells, as detected by HLA-A2-restricted CTL. According to the crystal structure of HLA-A2 (ref. 6), the critical residues for recognition by MA2.1 antibody, residues 62–66 of the $\alpha 1$ domain⁸, are in a region important for both T-cell recognition and peptide binding. Residue 63, moreover, has an interdomain contact with residue 171 of the $\alpha 2$ -helix⁹. We have previously reported that HLA-A2-restricted CTL specific for matrix protein of the influenza A virus recognize the synthetic peptide representing residues 58–68 of matrix protein (M58–68)¹⁰. Some of these CTL are also able to recognize synthetic peptide in association with HLA-Aw69, encoded by an allele closely related

to that of HLA-A2, but which does not bind MA2.1 antibody¹¹.

Surprisingly, when added directly to a cytotoxicity assay, the MA2.1 antibody greatly enhanced the recognition of peptide M58–68 on autologous target cells by peptide-specific HLA-A2/Aw69 cross-restricted CTL (Fig. 1a). This effect was consistently seen at antibody concentrations greater than $0.1 \mu\text{g ml}^{-1}$. On a target cell expressing HLA-Aw69 and HLA-B57, however, no enhancement of recognition by the same CTL was seen (Fig. 1b), even though the antibody still bound to both T cells and target cells. Another HLA-A2-specific monoclonal antibody, PA2.1 (ref. 12; cross-reactive with HLA-Aw69) did not show any shift in the dose-response curve for the recognition of peptides on target cells expressing either HLA-A2 or HLA-Aw69 (Fig. 1a, b). The slight inhibition of recognition shown in Fig. 1b with MA2.1 antibody could be due to effects on the T cell, because incubation of HLA-A2-restricted CTL in high concentrations of MA2.1 antibody can inhibit their ability to lyse target cells (data not shown). The enhancement is not unique to this peptide, as similar effects have been shown for the recognition of a different peptide, from the nucleoprotein of influenza B virus, recognized by specific HLA-A2-restricted CTL (data not shown; and P. Robbins, personal communication).

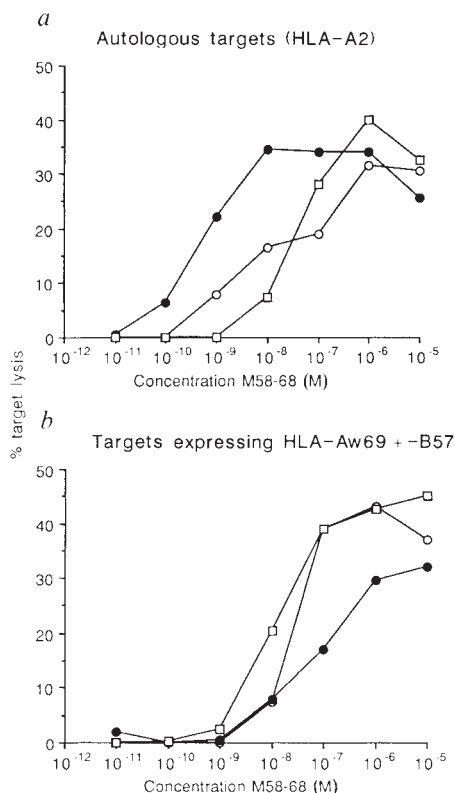


FIG. 1 The anti-HLA-A2 antibody, MA2.1, enhances recognition of peptide by matrix-specific HLA-A2/Aw69 cross-restricted CTL on HLA-A2—but not HLA-Aw69—expressing target cells. Effect of the presence of antibodies MA2.1 or PA2.1 on the dose-response curve of recognition of the peptide M58–68 of influenza matrix protein by CTL clone 1A10 from the donor JM (A2, 2; B15, 51; DR4, 4) on a, autologous target cells or b, a target cell MAL-NF (A1, w69; B35, 57; DR7, 10) which binds MA2.1 antibody on its B57 allele product. Both peptide and antibody ($1.25 \mu\text{g ml}^{-1}$) were added directly to the assay. Whereas MA2.1 antibody enhanced recognition of peptide by 10- to 100-fold on autologous cells, a small degree of inhibition was seen with the HLA-Aw69 expressing cells. Antibody PA2.1 did not produce a shift in dose-response curve in either case. Killer: target ratio (K:T)=1.8. Lysis in the absence of peptide, <2.3%. ○, No antibody; ●, MA2.1; □, PA2.1. METHODS. Influenza A virus matrix peptide M58–68 (Gly-Ile-Leu-Gly-Phe-Val-Phe-Thr-Leu-Thr-Val) was synthesized and donated by Dr J. Rothbard (ICRF, London). CTL clone 1A10 was derived as previously described¹¹. Purified monoclonal antibodies MA2.1 (ref. 3) and PA2.1 (ref. 12) were used, and were donated by Dr J. G. Bodmer (ICRF, London). ⁵¹Chromium-release cytotoxicity assay: Epstein-Barr virus-transformed lymphoblastoid cells (BCL), used as target cells, were labelled with ⁵¹Cr and added to the assay with CTL and diluted peptide and antibody, where appropriate, for 4 h in round-bottomed 96-well plates. Percentage target lysis was calculated from the formula $(E - M/D - M) \times 100$. E, experimental ⁵¹Cr release; M, release in presence of culture medium (always <18%); D, release by 5% Triton X-100.

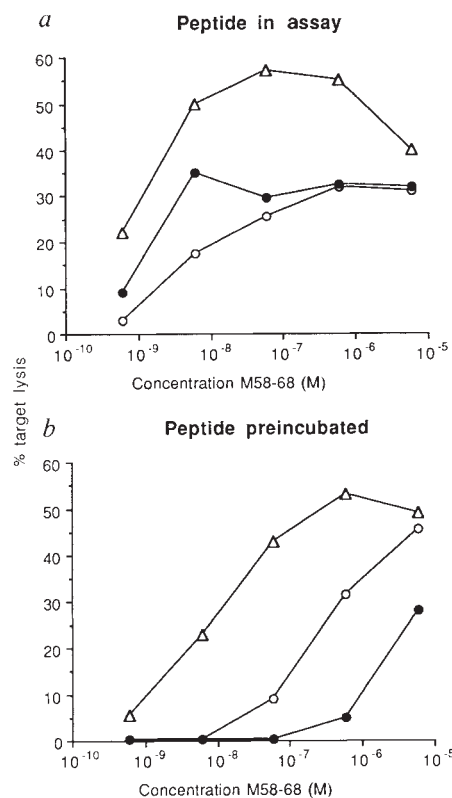


FIG. 2 Optimum enhancement of recognition by MA2.1 antibody occurred when target cells were exposed to antibody before or during exposure to peptide, and when MA2.1 antibody was not present in the assay. MA2.1 antibody inhibited recognition of target cells which had been pretreated with peptide. CTL clone 1A10-recognition of peptide M58–68 on autologous BCL target cells (K: T=3) was affected differently depending on the order of exposure of BCL to MA2.1 antibody or peptide. a, Peptide M58–68 titrated in the assay at the concentrations shown; target cells were either untreated (○) or pretreated with MA2.1 antibody ($1 \mu\text{g ml}^{-1}$) (△), MA2.1 antibody added directly to the assays (●). Lysis in the absence of peptide, <1%. b, Target cells were pretreated with peptide M58–68 at the concentrations shown. ○, No antibody; △, simultaneous pretreatment with MA2.1 antibody and peptide; ●, BCL pretreated with peptide with MA2.1 antibody added directly to assay. METHODS. As in Fig. 1, except that target cells pretreated with peptide were incubated, in dilution of peptide shown, for 1 h and then washed three times before inclusion in the assay.

These results show that the enhancement of peptide recognition by MA2.1 antibody is due to binding to the peptide-presenting HLA-A2 of the target cell rather than to HLA-A2 on the T cell, or to an irrelevant allele product (that is, non-presenting) on the target cell. To optimize the enhancement, various conditions of exposure to peptide and antibody were investigated. Figure 2a shows the results of experiments in which peptide was titrated in the CTL assay, and target cells were either pretreated with antibody, or antibody was added directly to the assay. Pretreatment with antibody enhanced peptide recognition ~100-fold with greatly increased lysis, whereas adding antibody to the assay only gave a modest increase. Figure 2b shows the results of experiments in which target cells were pretreated with peptide, and antibody was either added with the peptide or added directly to the assay. Whereas pretreatment with antibody again produced nearly 100-fold enhancement of peptide recognition, addition of MA2.1 antibody to the assay inhibited the recognition of the peptide-treated target cells, similar to the inhibition of recognition of virus-infected cells (refs 4 and 5, and data not shown). Thus, MA2.1 antibody has a dual action; first, it enhances peptide association at the level of HLA-A2 on the target cell; and second, it interferes with T-cell recognition when it is present in the assay.

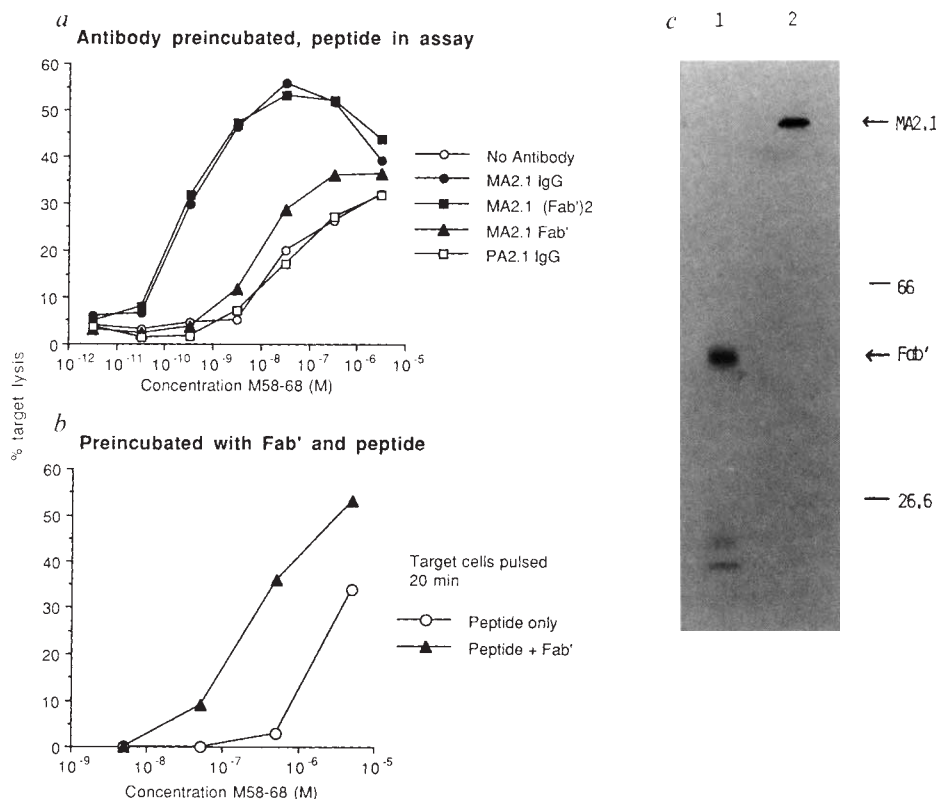
Because both MA2.1 and PA2.1 are IgG1 monoclonal antibodies¹³, and only MA2.1 antibody enhances peptide recognition, it is unlikely that the IgG1 Fc receptor mediates the effect. As antibody divalency could be important, either by crosslinking target cells and CTL, or by crosslinking HLA-A2 on the target-cell surface, the effects of F(ab')₂ and Fab' fragments of MA2.1

antibody were assessed. Target cells were pretreated with MA2.1 antibody or its fragments, using PA2.1 antibody as a control. CTL (Fig. 3a) were able to recognize peptide at 1,000-fold lower concentrations on HLA-A2-matched target cells pretreated with either whole MA2.1 antibody or its F(ab')₂ fragment relative to control. The enhancement seen with Fab' was much less, although consistent in five different experiments, and could be due to lower avidity of the monovalent fragment⁷. But a 10-fold enhancement of peptide recognition was seen with target cells incubated simultaneously with peptide and Fab' (Fig. 2b). This effect is unlikely to be due to contamination of the sample with divalent fragments, because SDS-PAGE of a radio-iodinated sample did not show any trace of either whole IgG or F(ab')₂ (Fig. 3c).

The simplest explanation of the data presented here is that the presence of MA2.1 antibody on a target cell enables HLA-A2 to bind peptide at the cell surface more efficiently. To test this hypothesis, the kinetics of target-cell sensitization with peptide were studied on untreated and MA2.1 antibody-pretreated cells. Targets were pulsed with 5 μ M peptide (Fig. 4a): in the absence of antibody, maximum lysis was not reached by 20 min of incubation with peptide. But when the target cell was saturated in antibody, a 30-s peptide pulse was sufficient to obtain almost maximal lysis. The lower concentration of 0.05 μ M peptide was insufficient to give any lysis in the absence of antibody, but with the antibody-treated cells, the lysis was as effective as it was at the 100-fold greater peptide concentration in the absence of antibody.

The rapidity of target sensitization in the presence of antibody

FIG. 3 Fragments of MA2.1 antibody are able to enhance recognition of peptide. Effect of MA2.1 antibody or its Fab' or F(ab')₂ fragments (1 μ g ml⁻¹; that is, equivalent numbers of combining sites for Fab' and F(ab')₂) on the dose-response curve of recognition of peptide M58-68 of the polyclonal CTL line JM 22 on HLA-A2-matched BCL target cells (A2, 11; B54, 0; DR4, 12). K:T=2. Lysis in the absence of peptide, <3.5%. a, BCL target cells were pretreated in antibody MA2.1 or PA2.1 or fragment (1 μ g ml⁻¹) as indicated and washed before inclusion in the assay. Both MA2.1 antibody and its F(ab')₂ fragment caused a 100- to 1,000-fold enhancement of recognition; the Fab' fragment enhanced recognition by ~10-fold. No enhancement was seen after preincubation with PA2.1. b, Recognition of autologous target cells by JM 22CTL was enhanced 10-fold (K:T=2) after pre-treatment of the target cells with a simultaneous pulse of peptide and Fab' (5 μ g ml⁻¹). Target cells were either pulsed with peptide alone at the concentrations shown, or with peptide with MA2.1 Fab', for 20 min, followed by three washes before addition to the assay. c, SDS-PAGE analysis of iodinated samples of the Fab' fragment of MA2.1 antibody (lane 1) or whole MA2.1 antibody (lane 2); 10% polyacrylamide gel, under nonreducing conditions. Relative molecular mass markers indicated ($\times 10^{-3}$). METHODS. JM 22 CTL were derived and maintained as previously described for other polyclonal matrix-specific lines¹⁰. F(ab')₂ and Fab' fragments of MA2.1 antibody were prepared from purified antibody by pepsin digestion, according to the method previously described for mouse IgG1 monoclonal antibodies¹³. Purity of fragments was established by SDS-PAGE (see c, and data not shown). Fragments were still able to bind HLA-A2 specifically as indicated by indirect fluorescence binding (data not shown). ⁵¹Cr release assay: As for Fig. 1 except for a; BCL target cells were pretreated in antibody or fragments for 1 h and then washed three times before inclusion in the assay. b, Pretreatment of



cells was for 20 min. SDS-PAGE of iodinated Fab' and whole antibody: Whole antibody or Fab' fragment (~10 μ g) was iodinated by addition of ¹²⁵I (0.5 μ Ci) in Chloramine-T (2.4 g ml⁻¹) for 2 min. The reaction was stopped with saturated tyrosine, and the iodinated fraction was separated on a PD-10 column. Each sample (10⁶ c.p.m.) were analysed on a 10% acrylamide gel by SDS-PAGE under nonreducing conditions. The gel was fixed, stained, dried and exposed to X-ray film.

makes an interaction at the cell surface a likely explanation for the enhancement of peptide recognition. This is supported by the fact that pretreatment of target cells with MA2.1 antibody before virus infection does not enhance the recognition of these cells (data not shown). Up-regulation of expression of class I molecules or HLA-A2 is an unlikely explanation, both because cycloheximide (an inhibitor of protein synthesis) has no effect on the enhancement, and antibody-binding studies have not shown an increase in cell surface HLA-A2 after MA2.1 antibody treatment (data not shown). The experiment represented in Fig. 1b rules out the crosslinking of T cell and target as a possible explanation.

The cell-surface effect could result from an increase in the number of peptide molecules bound in an immunogenic conformation, or from an increase in the total number of peptide molecules bound to HLA-A2. If, as predicted by the critical binding residues 62–66, MA2.1 antibody binds at the side of the groove in HLA-A2, the antibody could stabilize a conformation of A2 that allows freer exchange of peptide. As the T-cell receptor probably binds at a similar site, it is possible that the receptor could also facilitate antigenic peptide exchange. This

could be an important factor in the dissociation of the T-cell receptor-peptide-MHC complex once the T cell has been triggered.

These experiments demonstrate that peptide can rapidly associate with HLA-A2 at the surface of the cell in the presence of MA2.1. The antibody MA2.1 could be a useful tool in the preparation of a specific peptide with HLA-A2 for structural studies. □

Received 5 September; accepted 11 October 1989.

1. Townsend, A. R. M. *et al. Cell* **44**, 959–968 (1986).
2. Townsend, A. R. M. *et al. Nature* **340**, 443–448 (1989).
3. McMichael, A. J., Parham, P., Rust, N. & Brodsky, F. M. *Hum. Immun.* **1**, 121–129 (1980).
4. McMichael, A. J., Parham, P., Brodsky, F. M. & Pilch, J. R. *J. exp. Med.* **152**, 195s–203s (1980).
5. Dongworth, D. W. & McMichael, A. J. *Br. med. Bull.* **40**, 254–261 (1984).
6. Bjorkman, P. J. *et al. Nature* **329**, 512–518 (1987).
7. Ways, J. P. & Parham, P. *Biochem. J.* **216**, 423–432 (1983).
8. Santos-Aguado, J., Barbosa, J. A., Biro, P. A. & Strominger, J. L. *J. Immun.* **141**, 2811–2818 (1988).
9. Bjorkman, P. J. *et al. Nature* **329**, 506–512 (1987).
10. Gotch, F. M. *et al. Nature* **326**, 881–882 (1987).
11. Bodmer, H. C., Gotch, F. M. & McMichael, A. J. *Nature* **337**, 652–655 (1989).
12. Parham, P. & Bodmer, W. F. *Nature* **276**, 397–399 (1978).
13. Parham, P. *J. Immun.* **131**, 2895–2902 (1983).

ACKNOWLEDGEMENTS. We thank Dr P. Robbins for influenza B-specific CTL, Dr J. Rothbard for the peptides, Dr J. Bodmer for PA2.1 antibody and MAL-NF cells, CETUS corporation for recombinant interleukin-2 and Dr A. Townsend for discussion. H.C.B. is in receipt of an MRC Training Fellowship, and this work was supported by the MRC.

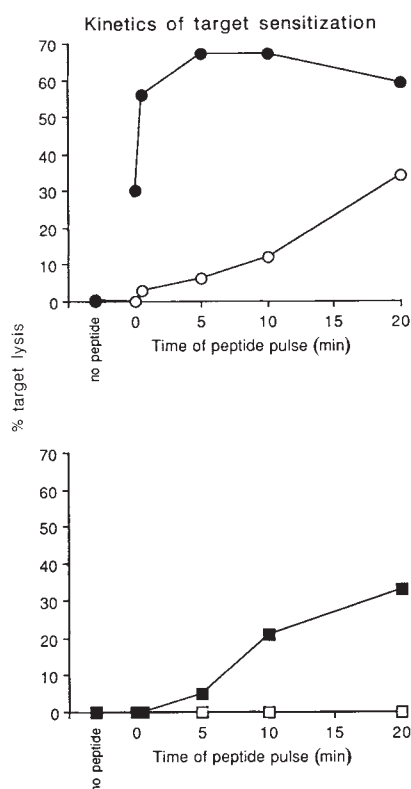


FIG. 4 Target sensitization with peptide occurs much more rapidly and at lower concentrations of peptide on target cells for target cells pretreated with MA2.1 antibody (●, ■) compared with untreated cells (○, □). Lysis by JM 22 CTL (K: T=2) of autologous JM BCL target cells which either had been pretreated with MA2.1 ($5 \mu\text{g ml}^{-1}$) or had not been treated, were pulsed with peptide M58–68 ($5 \mu\text{M}$ above; $0.05 \mu\text{M}$, below) for various time intervals from 30 s to 20 min.

METHODS. JM 22 CTL and ^{51}Cr -release assay as described in other figure legends, except for preparation of target cells, which were labelled with ^{51}Cr and either treated with MA2.1 ($5 \mu\text{g ml}^{-1}$) or not treated, for 1 h. Washed cells were then pulsed with peptide for the appropriate time interval followed by an immediate 100-fold dilution in serum-free medium at 4°C and then washed twice in RPMI-1640 at 4°C followed by a further wash in medium containing 10% FCS at room temperature. Cells were then diluted and added to the assay in the usual way. Time zero represents the lysis due to peptide which associated with the cells as a result of the 100-fold dilution of peptide M58–68 in the first wash at 4°C . Spontaneous release of ^{51}Cr in the absence of CTL was $<17\%$.

Insertion of cutinase gene into a wound pathogen enables it to infect intact host

M. B. Dickman*, G. K. Podila† & P. E. Kolattukudy†‡

* Department of Plant Pathology, University of Nebraska, Lincoln, Nebraska 68583, USA

† Ohio State Biotechnology Center, The Ohio State University, Columbus, Ohio 43210, USA

MANY phytopathogenic fungi must breach the intact plant cuticle to successfully invade and colonize their hosts. It has been suggested that cutinase, an extracellular enzyme secreted by many fungi, is essential for infection in some host–pathogen interactions. Chemical or immunological inhibition of cutinase protects the host from infection^{1,2}. We have cloned and characterized the cutinase complementary DNA and gene from *Fusarium solani* f. sp. *pisi*^{3,4}, a pea pathogen. A construct containing the cutinase coding region and extensive portions of the 5' and 3'-flanking regions from the *Fusarium* genome was transferred into another phytopathogenic Ascomycete, *Mycosphaerella* spp., a parasitic fungus that affects papaya fruits only if the fruit skin is mechanically breached before inoculation⁵. Here we describe the introduction of the *Fusarium* cutinase gene into *Mycosphaerella* to yield transformants in which this gene is inducible by cutin hydrolysate. We demonstrate that these transformants of the wound-requiring fungus have the capacity to infect intact papaya fruits, and that this infection can be prevented by antibodies against *Fusarium* cutinase.

Cutin, a biopolymer, is the chief structural component of plant cuticle⁶. Fungal contact with this polymer acts as a signal to induce cutinase production by the spores⁷. Chemically prepared cutin hydrolysate can also induce the cutinase gene, transcripts being detectable within fifteen minutes after the addition of inducer⁷. This induction allows the fungus to breach the cuticular barrier and penetrate into the host. *Mycosphaerella* spp. is unable to penetrate the outer defensive barrier of papaya fruits but can infect the fruits through wounds. To test whether this limitation of the pathogen could be overcome by the acquisi-

‡ To whom correspondence should be addressed.

tion of a functioning cutinase gene, we inserted this gene from *F. solani pisi* into *Mycosphaerella* spp. by a transformation procedure using whole fungal cells and lithium acetate treatment^{4,8}. The plasmid pU5-11 contains the cutinase structural gene and nearly 4 kilobases (kb) of 5' upstream sequence, plus 1 kb of 3' non-coding DNA⁴. *Mycosphaerella* was cotransformed with this plasmid and a plasmid containing the hygromycin-resistance gene^{4,9,10}; the transformants were selected for hygromycin resistance and screened for the cutinase gene by Southern blot analysis (Fig. 1). The four transformants shown in Fig. 1 were cut with *StyI*, an enzyme that does not cut within pU5-11. The integration events seem to be random. The transformants M-25 and M-101 (lanes 4 and 5) probably contain tandem duplications. Because signals of high molecular weight were present, it is likely that M-101 also has another separate site of integration.

We grew the transformants under inducing conditions¹¹ to see whether the cutinase gene that is induced by cutin monomers in *F. solani pisi*¹² is transcriptionally regulated in the *Mycosphaerella* transformants as well. Northern blots of transcripts from wild type and the four *Mycosphaerella* transformants were hybridized with radioactively labelled cutinase cDNA from *Fusarium* as probe (Fig. 1). An interval of 48 hours was chosen because in that time the level of transcription of the gene in *Fusarium solani pisi* reaches a maximum¹³. There was no hybridization with the wild-type *Mycosphaerella* RNA (lanes 1 and 2), showing that transcripts from this organism share no homology with *Fusarium* cutinase messenger RNA. Transformants M-12 (lane 4) and M-21 (lane 6) displayed similar patterns, with a unique transcript comparable in size to authentic *Fusarium* cutinase mRNA. Transformants M-25 and M-101 gave two and three separate transcripts, respectively. These multiple transcripts might have been due to differential processing or splicing of the mRNA; alternatively, insertion in different positions could lead to the production of transcripts of different

sizes in *Mycosphaerella*. A time course of induction of cutinase gene in transformant M-101 (Fig. 1, lanes 8–12) showed a low level of constitutive expression of the gene and maximal expression at 48 h after the addition of the inducer.

To investigate whether the transformants that produced transcripts containing the coding sequence for cutinase were translated, the culture supernatants from the induced cultures were subjected to western blot analysis, using as probe antibodies raised against *Fusarium solani pisi* cutinase. There was no cutinase detectable from the wild type even after 72 h of induction (Fig. 2). All four transformants produced cutinase within 72 h of addition of cutin hydrolysate. Within 24 h of adding inducer, the transformant M-101 was producing cutinase, which increased up to 72 h; For transformant M-25, cutinase levels were fairly high by 72 h, reaching 74% of the M-101 cutinase levels, although the induction was slower (data not shown). Transformants M-21 and M-12 yielded much less cutinase and this could only be found 72 h after the addition of the inducer.

Next we tested the proteins that were generated by the transformants and which cross-reacted with antibodies against *Fusarium* cutinase to see whether they showed enzyme activity. Cutinase activity correlated with the levels of the proteins detected by immunoblot analyses (Fig. 2): transformant M-101 gave the highest cutinase activity, followed by M-25, whereas supernatants from M-21 and M-12 were much less active. Thus, the levels of transcription, translation product and cutinase activity from the transformants were all in agreement.

We used a papaya fruit bioassay⁵ to investigate the pathogenicity of *Mycosphaerella* transformants expressing the inducible cutinase gene. Both wild-type *Mycosphaerella* and the transfor-

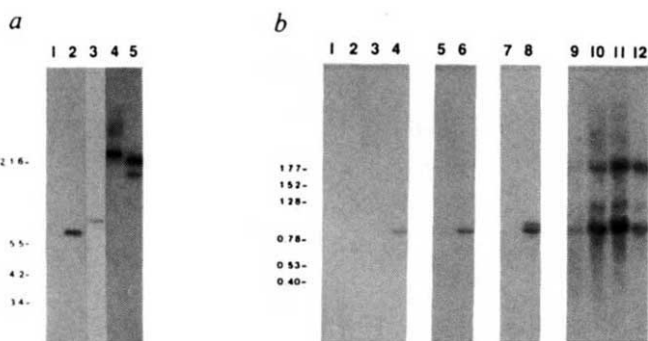


FIG. 1 *a*, Southern blot analysis of genomic DNA from wild-type *Mycosphaerella* spp. and transformants. Lane 1, wild type; lanes 2, 3, 4 and 5 represent transformants M-12, M-21, M-25 and M-101, respectively. Co-transformation was as described⁴, with 10 μ g of each plasmid yielding 2–4 stable transformants per μ g; about half of these showed mitotic stability and were analysed by Southern blotting. Genomic DNA from these transformants was digested with *StyI*, electrophoresed in 0.7% agarose, transferred to nitrocellulose and probed with nick-translated ³²P-labelled cDNA for cutinase from *F. solani pisi*³. Sizes in kilobases and migration of lambda *HindIII*-*EcoRI* fragments are shown on the left. *b*, Northern blots showing induced cutinase transcript levels in the transformants of *Mycosphaerella* spp. Lanes 1 and 2, wild-type *Mycosphaerella* uninduced and induced for 48 h respectively; lanes 3 and 4, 5 and 6, 7 and 8 represent transformants M-12, M-21, M-25 respectively: in each case the odd number is uninduced and the even number is induced for 48 h; lanes 9, 10, 11, and 12 represent transformant M-101 induced for 0, 24, 48 and 72 h, respectively. In each case glucose-grown fungi were induced with cutin hydrolysate (10 mg ml⁻¹) and RNA isolated with guanidium hydrochloride¹⁶ was separated on 1% formaldehyde agarose gels, transferred to nylon and hybridized to ³²P-labelled cutinase cDNA, as described in the legend to Fig. 1*a*. The total RNA levels loaded in each lane did not vary, as we confirmed by washing off the cutinase probe and reprobing the blot with ³²P-labelled actin cDNA.

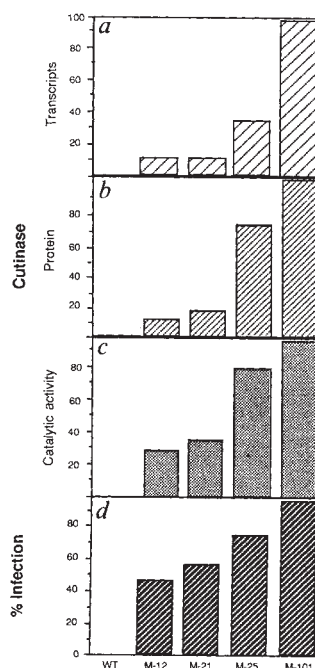


FIG. 2 Correlation of levels of cutinase transcripts (*a*), protein (*b*), enzymatic activity (*c*), with infectivity (*d*) of transformants of *Mycosphaerella* spp. The level of cutinase transcripts was measured by scanning the autoradiograph of the northern blot corresponding to a 48-hour period of induction, as described in the legend of Fig. 1*b*; the levels of cutinase protein produced were measured by scanning the western blot of culture supernatants from 72 h-induced cultures; the levels of enzymatic activity of cutinase in the culture supernatants were determined using *p*-nitrophenylbutyrate as the substrate¹¹ after 72 h induction; the per cent infection of the wild type and the transformants on intact papaya fruit was measured as described⁵. In all cases, data were normalized against values for transformant M-101 taken as 100. The cutinase activity and per cent infection for M-101 were 43.7 nmol min⁻¹ ml⁻¹ and 90%, respectively.

mants cause lesions when inoculated on wounds (data not shown). The wild type did not cause lesions when inoculated on intact fruits without wounds (Fig. 3), but the transformants expressing the *Fusarium* cutinase gene did. Frequency and size of lesions were the largest for transformant M-101; nearly all locations inoculated were infected. In fact, the incidence of infection and the severity of the lesions caused by this transformant were essentially identical to those observed with *Colletotrichum gloeosporioides*, an authentic pathogen of intact papaya fruits⁵. The extent to which the transformants become parasites on papaya fruit therefore correlates extremely well with the expression of *Fusarium* cutinase gene (Fig. 2).

To test whether infection of intact fruits by the transformant was due to cutinase, rabbit antibodies against the cutinase from *F. solani pisi* were included with the spore suspension of the M-101 transformant, which produced the most cutinase and the most severe lesion. This antibody preparation, which inhibits *Fusarium* cutinase, prevented lesion formation by transformant M-101 (Fig. 3). Thus, it is clear that *Fusarium* cutinase activity produced by the *Mycosphaerella* transformant enabled it to penetrate the intact fruit and infect it.

We have shown that by inserting a gene encoding a cuticle-degrading enzyme from a pathogenic organism into one that infects only through wounds, the latter can now infect intact plant organs without a wound. As *Mycosphaerella* spp. is able to infect wounded tissue, then this fungus must have most of the genetic machinery required for parasitism. Apparently, however, it cannot penetrate the host's first line of physical defence, the cuticle. Overcoming this limitation by gene insertion allows this organism to infect intact papaya fruits.

Regulation of the *Fusarium* cutinase gene is unchanged when inserted into *Mycosphaerella*: on contact of the spore with the plant surface, gene expression is triggered by cutin monomers released by the small amount of cutinase carried by the spore. The transformants that efficiently expressed cutinase showed

low levels of constitutive expression (Fig. 1, lane 9; also confirmed by Western blot analysis, data not shown). The resulting level of cutinase would release monomers from papaya cutin. Dihydroxyhexadecanoic acid is a most effective inducer of the cutinase gene in *Fusarium*^{7,12} and is the most abundant monomer of papaya cutin¹⁴, so the appropriate inducer is available to the transformants when placed on the fruit surface, enabling them to use their regulated expression of cutinase to gain entry into the fruits and infect them. Activation of the cutinase gene in isolated nuclei from *F. solani pisi* by the unique cutin monomer has been demonstrated¹². This selective initiation of transcription of the cutinase gene required a protein factor from the *Fusarium* extract. Apparently *Mycosphaerella* spp. possesses this protein or a functional analogue, as it can use cutin hydrolysate to activate the *Fusarium* cutinase gene in a similar way. This is not surprising, as similar target proteins that regulate transcription are present in many eukaryotes¹⁵. By understanding the regulation of genes essential for pathogenesis, it might be possible to interfere with their expression and hence pathogenesis. □

Received 26 July; accepted 9 October 1989.

1. Maiti, I. & Kolattukudy, P. E. *Science* **205**, 507-508 (1979).
2. Kolattukudy, P. E. *A. Rev. Phytopath.* **23**, 223-250 (1985).
3. Soliday, C. L., Flurkey, W. H., Okita, T. W. & Kolattukudy, P. E. *Proc. natn. Acad. Sci. U.S.A.* **81**, 3939-3943 (1986).
4. Soliday, C. L., Dickman, M. B. & Kolattukudy, P. E. *J. Bact.* **171**, 1942-1951 (1989).
5. Dickman, M. B., Patil, S. S. & Kolattukudy, P. E. *Physiol. Plant Pathol.* **20**, 333-347 (1982).
6. Kolattukudy, P. E. *Science* **208**, 990-1000 (1980).
7. Woloshuk, C. P. & Kolattukudy, P. E. *Proc. natn. Acad. Sci. U.S.A.* **83**, 1704-1708 (1986).
8. Dhawale, S. S., Paletta, J. V. & Marzluf, G. A. *Curr. Genet.* **8**, 70-79 (1984).
9. Kaster, R. K., Burgett, S. G. & Ingolia, T. D. *Curr. Genet.* **8**, 353-358 (1984).
10. Turgeon, B. G., Garber, R. C. & Yoder, O. C. *Molec. cell. Biol.* **7**, 3297-3305 (1987).
11. Lin, T. S. & Kolattukudy, P. E. *J. Bact.* **133**, 942-951 (1978).
12. Podila, G. K., Dickman, M. B. & Kolattukudy, P. E. *Science* **242**, 922-925 (1988).
13. Kolattukudy, P. E., Podila, G. K. & Mohan, R. *Genome* **31**, 342-349 (1989).
14. Walton, T. J. & Kolattukudy, P. E. *Biochemistry* **11**, 1885-1897 (1972).
15. Ptashne, M. *Sci. Am.* **260**, 40-47 (1989).
16. Cathala, G. *et al. DNA* **2**, 329-335 (1983).

ACKNOWLEDGEMENTS. This work was supported by the National Science Foundation.

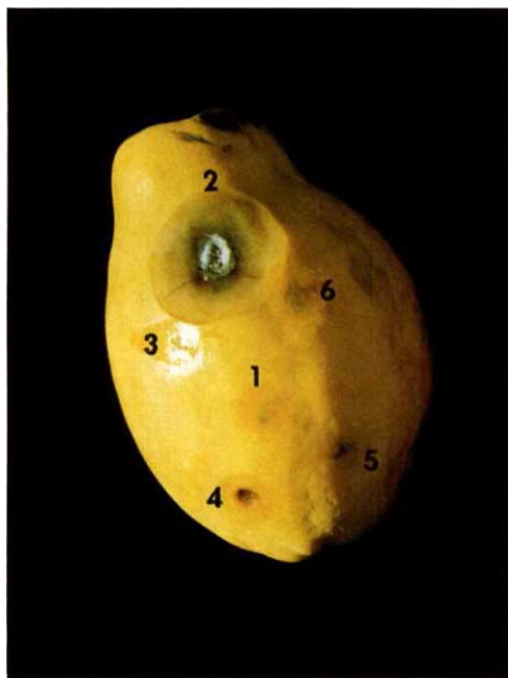


FIG. 3 Lesion formation on intact papaya fruits by *Mycosphaerella* transformants into which *Fusarium* cutinase gene was introduced. (1) Wild type *Mycosphaerella* spp.; (2) transformant M-101; (3-5) transformants M-12, M-25 and M-21, respectively; (6) transformant M-101 with rabbit antibodies prepared against *Fusarium* cutinase. The picture was taken 72 h after keeping the intact papaya fruits with the spore suspension (5×10^4 spores in $50 \mu\text{l}$) on it at 26°C in a 100% humid chamber as before⁵.

Molecular basis of lipid transfer protein deficiency in a family with increased high-density lipoproteins

Maryanne L. Brown*, Akihiro Inazu*†, Charles B. Hesler*, Luis B. Agellon*, Chris Mann*, Mary E. Whitlock*, Yves L. Marcel‡, Ross W. Milne‡, Junji Koizumi†, Hiroshi Mabuchi†, Ryoyu Takeda† & Alan R. Tall*§

* Department of Medicine, Columbia University College of Physicians and Surgeons, New York, New York 10032, USA

† The Second Department of Internal Medicine, Kanazawa University School of Medicine, Kanazawa 920, Japan

‡ Clinical Research Institute of Montreal, Pine Avenue West, Montreal H2W 1R7, Canada

PLASMA high density lipoproteins (HDL) are a negative risk factor for atherosclerosis. Increased HDL is sometimes clustered in families, but a genetic basis has never been clearly documented¹. The plasma cholesteryl ester transfer protein (CETP) catalyses the transfer of cholesteryl ester from HDL to other lipoproteins and therefore might influence HDL levels². Using monoclonal antibodies, we show that CETP is absent in two Japanese siblings who have markedly increased and enlarged HDL. Furthermore,

§ To whom correspondence should be addressed.

they are homozygous for a point mutation in the 5'-splice donor site of intron 14 of the gene for CETP, a change that is incompatible with normal splicing of pre-messenger RNA³. The results indicate that the family has an inherited deficiency of CETP due to a gene splicing defect, and illustrate the key role that CETP has in human HDL metabolism.

A Japanese subject has recently been described with markedly increased HDL and low levels of cholesteryl ester transfer activity⁴. Low activity could reflect a deficiency of CETP, the presence of an inhibitor⁵, or a substrate effect (high HDL)⁶. Our goals were to assess blood and plasma samples from the Japanese subject and his family for the presence of the CETP of relative molecular mass (M_r 74,000 (74K))⁷, using two different CETP monoclonal antibodies^{8,9}, and also to analyse the defect at the DNA level using the sequence of the wild-type CETP gene (L.B.A. *et al.*, manuscript in preparation).

CETP in plasma was immunoaffinity-purified using an anti-CETP monoclonal IgG (TP2) column⁸ and the retained fraction was analysed by western immunoblotting of SDS-polyacrylamide gels (Fig. 1). In the control subjects, in T.Y. (proband's spouse) and in the four children (Y.Y., K.C., M.D., K.Y.), CETP ran at its characteristic position⁷. Remarkably, CETP was not detected in the plasma of the proband (S.Y.) or of his sister (K.H.) (Fig. 1), even when 300 μ l of S.Y. plasma was analysed (S.Y. $\times 3$) (Fig. 1). Also, when equal amounts of control and S.Y. plasma were mixed, there was no diminution of the CETP signal following immunoaffinity chromatography (Fig. 1; con + S.Y.).

To quantitate CETP more precisely, plasma samples were analysed by radioimmunoassay using ¹²⁵I-labelled TP2. In this competitive displacement assay¹⁰, the plasma of S.Y. and K.H. did not displace the standard CETP from the solid phase, which is consistent with no CETP being present (Table 1). The amount of plasma CETP from the children ranged from 38 to 56% (mean 49%) of their mother's value (Table 1), which is consistent with their being obligate heterozygotes for inherited CETP deficiency, with codominant expression determining CETP levels.

In a further attempt to detect CETP in the S.Y. plasma, it was electrophoresed in native polyacrylamide gradient gels and

either stained with Sudan Black to show the lipoproteins (Fig. 2a) or transferred onto nitrocellulose and probed with a monoclonal antibody that recognizes a different epitope, ¹²⁵I-labelled TP6 (Fig. 2b). In the lipid-stained preparations, normal HDL appears as a set of diffusely stained bands corresponding to particles of diameter between 8 to 12 nm (Fig. 2a; controls 1-4), and low-density lipoproteins (LDL) run as an intense band near the top of the gel. The proband (S.Y.) shows markedly increased and enlarged HDL, extending into the size region of LDL (Fig. 2a). Similar results were obtained with his sister (K.H.) (data not shown). Two of the children (K.C. and Y.Y.) showed a distinct increase in staining of the larger HDL subclass, and all showed an increase in HDL-2 (10-12 nm) relative to the smaller HDL-3 (8-10 nm). Thus, the HDL₂/HDL₃ ratio was above normal in the children (Table 1).

Western immunoblot analysis of plasma run on native polyacrylamide gradient gels and probed directly with ¹²⁵I-labelled TP6, showed that the CETP was present in a diffuse band co-migrating with HDL-3, with a similar distribution in control samples and in samples from T.Y. and the four children (Fig. 2b). But, no CETP could be detected in the plasma of either

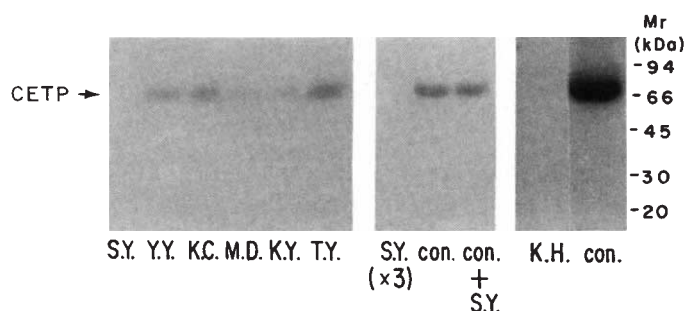


FIG. 1 Western immunoblot analysis of the anti-CETP immunoaffinity-retained fraction of plasma. Results are shown for the proband (S.Y.), his spouse (T.Y.), their four children (Y.Y., K.C., M.D., K.Y.), normolipidemic control subjects (con) and also for S.Y.'s sister (K.H.). The 'S.Y. ($\times 3$)' incubation contained 300 μ l S.Y. plasma and the 'con + S.Y.' incubation consisted of 100 μ l each of control and S.Y. plasma; all other incubations contained 100 μ l plasma from the subject indicated. Autoradiographs (3-day exposure, except for K.H. and control, which is a 6-day exposure) of these blots are shown. The position of the 74K CETP is indicated as well as the positions of marker proteins (Pharmacia).

METHODS. Plasma samples containing 1 mM PMSF were subjected to immunoaffinity chromatography on an anti-CETP monoclonal IgG (TP2) column as previously described⁸. Recovery of CETP in the retained fraction under these conditions was $\sim 80\%$. For western immunoblots, the retained fractions were subjected to SDS-PAGE, transferred to nitrocellulose and incubated with ¹²⁵I-labelled (Iodobeads-Pierce) anti-CETP monoclonal antibody TP2 (5×10^6 c.p.m./per ml).

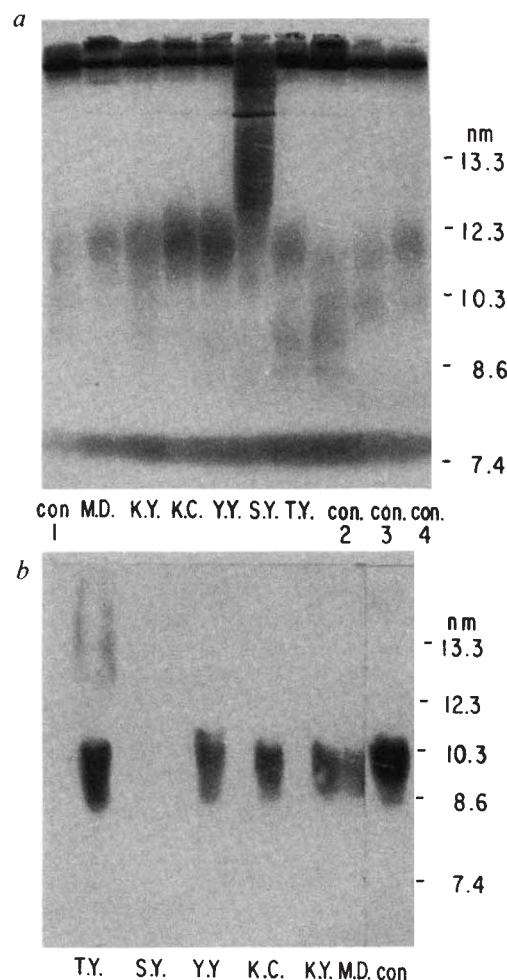


FIG. 2 Native, gradient polyacrylamide gel *a*, stained for lipid with Sudan Black, or *b*, western immunoblotted using anti-CETP monoclonal antibody TP6. *a*, The lipoprotein profile of the S.Y. family (see legend to Fig. 1) and 4 different normolipidemic controls was visualized by Sudan Black staining of a 4-30% gradient acrylamide gel (Pharmacia). Each lane represents 30 μ l plasma electrophoresed for 18 h in Tris-borate buffer, pH 8. *b*, Western immunoblot analysis of 30 μ l plasma from S.Y. and other family members after reaction with ¹²⁵I-labelled anti-CETP monoclonal antibody TP6. The hydrodynamic diameters (in nm) were estimated from the mobility of Pharmacia protein standards, which are shown on the right.

TABLE 1 HDL and cholesteryl ester transfer protein in the SY family

		Age/Sex	HDL-chol (mg dl ⁻¹)	APOA-I* (mg dl ⁻¹)	HDL2/HDL3†	CETP Activity (% control)	CETP‡ Mass (µg ml ⁻¹)
Proband	S.Y.	64 M	248	239	—	0	0
Siblings	K.H.	60 F	175	147	—	0	0
	T.A.	57 F	55	129	—	35	—
	Y.K.	47 M	63	—	—	105	—
Spouse	T.Y.	58 F	64	—	1.3	110 ± 10	2.5 ± 0.6
Children	M.D.	35 F	52	98	2.3	60 ± 12	1.0 ± 0.2
	K.C.	33 F	65	132	4.1	62 ± 5	1.3 ± 0.3
	Y.Y.	31 F	77	132	5.2	80 ± 14	1.4 ± 0.3
	K.Y.	28 M	52	98	4.1	45 ± 16	1.2 ± 0.4
Normolipemic (mean ± s.d.)			48 ± 13	121 ± 22	0.99 ± 0.20		2.0 ± 0.5
Controls							
	range		35–65	83–155	0.7–1.7		1.2–2.7
	N		19	19	15		12

* Determined by radial immunodiffusion.

† Determined by scanning Sudan-stained native gels (mean of duplicate determination).

‡ Determined by radioimmunoassay (mean ± s.d. of 3 separate assays).

S.Y. (Fig. 2b) or his sister (data not shown). Thus, analysis of plasma with two different antibodies against CETP (TP2 and TP6) (ref. 9) indicated an absence of CETP in both S.Y. and K.H.

This unusual lipoprotein phenotype raised the question of whether CETP deficiency was a secondary phenomenon or the result of an inherited defect of the gene for CETP. Preliminary Southern blot analysis using the CETP complementary DNA¹¹, revealed no major insertions or deletions of the CETP gene in the proband (results not shown). Therefore, the proband's genomic DNA (S.Y.) was used as a substrate for amplification of the 16 exons and exon/intron junctions of the CETP gene by the polymerase chain reaction (PCR)¹², in which we used intronic oligonucleotide primers positioned 50 to 100 nucleotides from the junctions. The amplified products were subcloned and sequenced¹³. The sequenced exons and exon/intron junctions encompassing the entire CETP gene (including the polyadenylation region and the translation initiation codon) were conserved when compared with the wild-type sequences (L.B.A. *et al.*, manuscript in preparation), with one exception. At the 5' splice donor of intron 14 (position +1) there was a G → A change altering the strictly conserved G-T intron splice donor¹⁴ to A-T. This change was identified in three different subclones derived from separate PCR amplifications of the proband's DNA.

The G → A mutation was confirmed and familial inheritance documented by direct sequencing of amplified DNA from all available family members (Fig. 3). The amplified exon 13,14 PCR product was directly sequenced using a primer originating within exon 14. Figure 3 shows the DNA sequencing ladder in the region of the exon 14/intron 14 junction. Columns 1 and 2 represent the homozygous CETP-deficient individuals (K.H. and S.Y.) and the arrow indicates the A at the first position (+1) of intron 14. Columns 3, 4, 5 and 6 represent the children of the proband and they are heterozygotes with both an A and a G at this position of the sequencing ladder. The mother (T.Y.) of the heterozygotes is represented under column 7 with the normal G at this site. The first nucleotide of a normal intron is never an A (ref. 15). Mutations of the strictly conserved G at this position prevent normal splicing¹⁶. The position +1 G is critical for correct splicing as it is involved in binding U1 small nuclear ribonucleoprotein particles (U1 snRNPs) and is the residue that undergoes cleavage in the spliceosome and lariat formation¹⁷. A similar G → A mutation in the conserved splice donor site in intron 1 or intron 2 results in two forms of β -thalassaemia. Both forms have abnormal β -globin mRNA and no β -globin protein^{18,19}. In some cases, splice site mutations lead to activation of nearby cryptic splice sites²⁰ or use of an alternative normal splice site (exon skipping)²¹, resulting in a truncated transcript and/or abnormal protein. The epitope of

TP2 is at the C-terminus of CETP⁹, corresponding to a region of the CETP gene downstream of the mutation. But CETP is also undetectable with the TP6 antibody (Fig. 2b), whose epitope maps between amino acids 261 to 367 of CETP⁹. This region of CETP is encoded by exons proximal to the site of the mutation. Thus, even if an alternative splice site is used in the mutant CETP gene, no CETP polypeptide is found in plasma. The dinucleotide ^mCpG is under-represented²² in the genome, but contributes to most (>35%) mutations characterized by single-base changes²³. We speculate that the CETP deficiency of S.Y. and K.H. might have involved a hot-spot deamination on the non-coding strand.

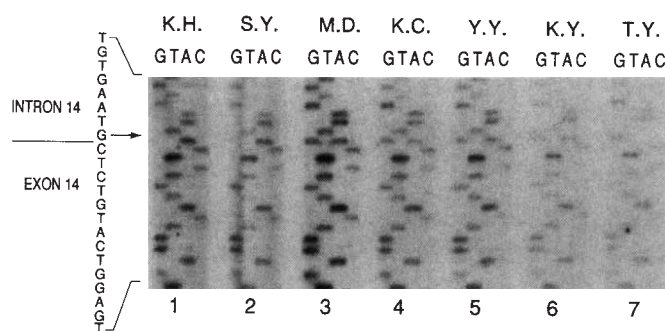


FIG. 3 DNA sequence ladder generated from the amplified products at the exon 14/intron 14 junction of the CETP gene in a family with CETP deficiency. Columns 1 and 2 (K.H. and S.Y.) of the autoradiograph (96 h exposure) represent the sequence of the homozygous CETP-deficient individuals. The arrow indicates the A residue observed in both K.H. and S.Y. Columns 3, 4, 5 and 6 (M.D., K.C., Y.Y. and K.Y.) show the heterozygotes with both an A and a G at this position of the sequencing ladder. The mother (T.Y.) of the heterozygotes is represented under column 7 with the normal G at this site.

METHODS. Genomic DNA was isolated²⁴ from 10 ml blood and used as a substrate for amplification by PCR¹². Polymerase chain reactions were carried out in a total volume of 100 µl under mineral oil, and contained 3 µg genomic DNA, 100 pmol primer A (5'-CCTGAGCTATGAGACAAAAG-3', located 92 nucleotides 5' of exon 13 of the gene for CETP), 100 pmol primer B (5'-AAAAGGTGAAATGGGAAGCT-3' located 43 nucleotides 3' of exon 14 of the gene for CETP), 2.5 U *Taq* polymerase (Cetus), 200 µM each dNTP, 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1.5 mM MgCl₂, 0.01% gelatin. The amplified products (30 cycles at 94°C for 0.5 min, 55°C for 1 min, 74°C for 3 min) were purified by Centricon-30 (Amicon) filtration and sequenced with a ³²P-labelled exon 14 primer (5'-CCAGAGCTTCCTGCAGTCAA-3' located 56 nucleotides 5' of the exon 14/intron 14 boundary) according to the Higuchi modifications of the 'Sequenase' USB protocol²⁵. The sequencing reactions were heat-denatured and resolved on a 6% polyacrylamide gel containing 7 M urea²⁶.

Our chief finding, therefore, is that an absence of CETP results from a single base change in the conserved splice donor site in a subject (S.Y.) and his sister (K.H.), both of whom have markedly increased and enlarged HDL. Furthermore, there were reduced levels of CETP in the children, obligate heterozygotes who carry the defect. They also have about half the normal levels of CETP and moderately increased or enlarged HDL. The magnitude of the increase in HDL (4–6 times normal) in S.Y. and his sister K.H. is surprising, and indicates that CETP is of major importance in HDL catabolism, influencing the size, composition and quantity of HDL. The mechanism leading to the increase in the HDL protein apoA-I in S.Y. and K.H. (Table 1) is unclear, but could result from slower catabolism of their

greatly enlarged HDL particles. It is also notable that low-density lipoproteins (LDL) were decreased in S.Y. and K.H.⁴, which is shown by a lack of Sudan Black staining in the LDL region (Fig. 2a). This finding indicates that the normal formation of LDL requires CETP. There are no obvious clinical manifestations of CETP deficiency in S.Y., his sister or the children⁴. The lipoprotein profile associated with human CETP deficiency (that is, high HDL and low LDL) has low atherogenic potential, raising the possibility that CETP inhibitors could be used as anti-atherogenic drugs. Indeed, the lipoprotein profile found in human CETP deficiency resembles that of species that lack CETP activity in their plasma (rats, dogs and pigs) and which are relatively resistant to atherosclerosis². □

Received 2 August; accepted 5 October 1989.

1. Breslow, J. L. in *Metabolic Basis of Inherited Disease* 6th ed., 1251–1266 (McGraw-Hill, New York, 1989).
2. Tall, A. R. *J. Lipid Res.* **27**, 361–367 (1986).
3. Padgett, R. A., Grabowski, P. J., Konarska, M. M., Seilela, S. & Sharp, P. A. *Rev. Biochem.* **55**, 1119–1150 (1986).
4. Koizumi, J. et al. *Atherosclerosis* **58**, 175–186 (1985).
5. Morton, R. E. & Zilversmit, D. B. *J. Biol. Chem.* **256**, 11992–11995 (1981).
6. Barter, P. J. & Jones, M. E. *J. Lipid Res.* **21**, 238–249 (1980).
7. Hesler, C. B., Swenson, T. L. & Tall, A. R. *J. Biol. Chem.* **262**, 2275–2282 (1987).
8. Hesler, C. B. et al. *J. Biol. Chem.* **263**, 5020–5023 (1988).
9. Swenson, T. L. et al. *J. Biol. Chem.* **264**, 14318–14326 (1989).
10. Marcel, Y. L. et al. *Arteriosclerosis* **8**, 593a (1988).
11. Drayna, D. et al. *Nature* **327**, 632–634 (1987).
12. Saiki, R. et al. *Science* **239**, 487–491 (1988).

13. Hattori, M. & Sakaki, Y. *Analyt. Biochem.* **152**, 232–238 (1986).
14. Mount, S. M. *Nucleic Acids Res.* **10**, 459–472 (1982).
15. Green, M. R. *A. Rev. Genet.* **20**, 671–708 (1986).
16. Aebi, M., Hornig, H., Padgett, R. A., Reiser, J. & Weissman, C. *Cell* **47**, 555–565 (1986).
17. Lamond, A. I., Konarska, M. M., Grabowski, P. J. & Sharp, P. A. *Proc. natn. Acad. Sci. U.S.A.* **85**, 411–415 (1988).
18. Treisman, R., Orkin, S. H. & Maniatis, T. *Nature* **302**, 591–596 (1983).
19. Baird, M. et al. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4218–4221 (1981).
20. Wieringa, B., Meyer, F., Reiser, J. & Weissmann, C. *Nature* **301**, 38–43 (1983).
21. Mitchell, P. J., Urlaub, G. & Chasin, L. *Molec. Cell Biol.* **6**, 1926–1935 (1986).
22. Bird, A. P. *Nucleic Acids Res.* **8**, 1499–1504 (1980).
23. Coulondre, C., Miller, J. H., Farabaugh, P. J. & Gilbert, W. *Nature* **274**, 775–780 (1978).
24. Kunkel, L. M. et al. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1245–1249 (1977).
25. Higuchi, R., von Beroldingen, C. H., Sensabaugh, G. H. & Erlich, H. A. *Nature* **332**, 543–545 (1988).
26. Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 449–560 (1980).

Demonstration by genetic suppression of interaction of GroE products with many proteins

Tina K. Van Dyk, Anthony A. Gatenby
& Robert A. LaRossa*

Central Research and Development Department, E. I. du Pont de Nemours and Company, Experimental Station, PO Box 80328, Wilmington, Delaware 19880-0328, USA

THE way in which proteins attain and maintain their final form is of fundamental importance. Recent work has focused on the role of a set of ubiquitous proteins, termed chaperonins¹, in the assembly of phage² and multisubunit proteins^{3,4}. The range of chaperonin action is unknown; they could interact with most cellular polypeptides or have a limited subset of protein partners. Included in the chaperonin family is the essential heat-shock regulated *Escherichia coli* *groEL* gene product^{1,5,6}. Overexpression of the *groE* operon in *E. coli* causes enhanced assembly of heterologously expressed ribulose biphosphate carboxylase subunits³ and suppresses the heat-sensitive mutant phenotype of several *dnaA* alleles^{7,8}. It has been inferred that suppression of heat-sensitive mutations is confined to *dnaA* alleles and that this confinement could reflect an interaction between the *groE* operon products and a *dnaA* protein aggregate at the replication origin^{7,9}. We now report that multiple copies of the *groE* operon suppress mutations in genes encoding several diverse proteins. Our data indicate a general role for the *groE* operon products, the GroEL and GroES proteins¹⁰, in the folding–assembly pathways of many proteins.

It has been proposed that complexes of enzymes involved in sequential metabolic reactions are common in cells¹¹. To investigate if enzymes of a biosynthetic pathway are part of a weakly

associated *in vivo* supramolecular structure, we isolated heat-sensitive mutations in the major *ilv* operon of *Salmonella typhimurium* (D. Smulski, T.K.V.D. and R.A.L., unpublished results) that confer a conditional growth requirement for isoleucine and/or valine. This was achieved by using a localized hydroxylamine mutagenesis protocol¹². It was expected that overexpression of the *groE* operon products might be able to correct growth defects by forcing either the folding or assembly of recalcitrant mutant polypeptides. The auxotrophic requirements of one mutation in *ilvGM* and two mutations in *ilvE*, encoding the multimeric enzymes acetolactate synthase II and transaminase B, respectively, were suppressed by the multicopy plasmid bearing the *groE* operon (Table 1). These data indicated that mutations in genes encoding biosynthetic enzymes can be suppressed by overexpression of the *groE* operon products.

We then investigated whether overexpression of the *groE* operon products could suppress mutations in the histidine (*his*) operon of *S. typhimurium* (Table 1). Heat-sensitive^{13,14}, cold-sensitive¹⁴ and temperature-independent¹³ auxotrophic mis-sense mutations in the *his* operon were examined. Many, but not all, heat-sensitive *hisD*, *hisC* and *hisB* alleles were suppressed by *groE* over-expression (Table 1); these genes encode, respectively, histidinol dehydrogenase, histidinol-phosphate aminotransferase and the bifunctional protein imidazoleglycerol phosphate dehydratase—histidinol phosphate phosphatase¹⁵. Suppression of cold-sensitive and temperature-independent alleles has so far not been observed in the *his* system (Table 1).

We quantitated the suppression of the *hisD6592* allele by *groE* overexpression with a sensitive radiochemical assay¹⁶, performed at 37 °C, of histidinol dehydrogenase activity present in toluene-permeabilized cells grown at the restrictive temperature of 37 °C in minimal medium E (ref. 17) supplemented with excess (0.64 mM) L-histidine. Equivalent repressed levels of histidine biosynthetic enzymes are produced in both auxotrophs and prototrophs under these conditions¹⁵. Specific activities thus reflect differences unrelated to transcription. A strain containing the *hisD6592* allele and the control plasmid pTG10, had undetectable enzyme activity (<0.5% of wild-type specific activity) whereas a strain containing the same *hisD* allele and the plasmid overexpressing *groESL* (ref. 3) contained 3%

* To whom correspondence should be addressed.

TABLE 1 Suppression of missense alleles

Alleles	Type	Restrictive temperature (°C)	Suppression
<i>S. typhimurium</i>			
<i>hisF8611*</i> , <i>hisF8612</i> , <i>hisF8613</i> ,	CS	18	—
<i>hisH8606</i>	HS	42	+
<i>hisB8435</i>			
<i>hisC6595</i> , <i>hisD141</i> , <i>hisD6585</i> ,			
<i>hisD6586</i> , <i>hisD6589</i> , <i>hisD6591</i> ,	HS	37	+
<i>hisD6592</i> , <i>hisD6593</i> , <i>hisC15</i>	HS	30	+
<i>hisC163</i>	HS	42	—
<i>hisC6596</i> , <i>hisC6598</i> , <i>hisC6599</i>	HS	39	—
<i>hisA8437</i> , <i>hisA8438</i> , <i>hisB8434</i> ,	HS	37	—
<i>hisC6594</i>	HS	30	—
<i>hisB59</i> , <i>hisB59</i>	HS	37	—
<i>hisG6582</i> , <i>hisG6583</i>	HS	30	—
<i>hisA338</i> , <i>hisB14</i> , <i>hisB20</i> ,			
<i>hisB53</i> , <i>hisC201</i> , <i>hisC899</i> ,	TI	NA	—
<i>hisD362</i> , <i>hisG46</i> , <i>hisG204</i> ,			
<i>hisI352</i>	HS	37	+
<i>ilvE3149†</i> , <i>ilvE3150</i>	HS	30	+
<i>ilvGM3139</i>	HS	42	—
<i>ilvGM3136</i> , <i>ilvGM3142</i> , <i>ilvGM3143</i> ,	HS	37	—
<i>ilvGM3147</i> , <i>ilvGM3148</i>	HS	30	—
<i>ilvGM3137</i>			
<i>ilvGM3135</i> , <i>ilvGM3138</i> , <i>ilvGM3140</i> ,	HS		
<i>ilvGM3141</i>			
<i>S. typhimurium</i> phage P22			
<i>9tsfU3‡</i> , <i>9tsfU8§</i> , <i>9tsfU34‡</i> ,	HS	39	+
<i>9tsfH304§</i>	HS	39	—
<i>9tsfU38§</i> , <i>9tsfN39§</i> , <i>9tsfU40§</i>	HS	37	—
<i>9tsfUT51</i> , <i>9tsu334§</i> , <i>9tsfH302§</i>	HS	30	—
<i>9tsfUT51§</i>			
<i>E. coli</i>			
<i>secA51 </i> , <i>secY24¶</i> ,	HS	42	+
<i>alaS4#</i> , <i>glnS1#</i> , <i>leuS31#</i> ,	HS	42	—
<i>metG83#</i> , <i>serS14#</i>			

Plasmids pTG10 (control)³ and pGroESL (*groES*⁺ *groEL*⁺)³ were transformed into a restriction-deficient strain of *S. typhimurium* LT2, TR5878 (ref. 28) selecting for chloramphenicol resistance. Phage P22-mediated generalized transduction^{17,29} was used to transfer the plasmid to the various *S. typhimurium* mutants. Growth of the *S. typhimurium* recombinants was tested at 18, 30, 37, 39 and 42 °C on minimal medium E supplemented with 0.4% glucose¹⁷. Introduction of pTG10 did not allow growth of any mutant at the restrictive temperature. (+), Strain containing plasmid pGroESL grew at temperatures restrictive for the growth of the otherwise isogenic plasmid pTG10 containing strain. Suppression of P22 mutations was tested by assaying plaque formation on lawns of *S. typhimurium* DB7136 (ref. 30) derivatives containing either plasmids pTG10 or pGroESL incubated at the restrictive temperature. Transformation was used to introduce the plasmids into *E. coli* strains. With the exception of *secY24*, suppression tests of all *E. coli* mutations was scored on chloramphenicol-supplemented LB plates¹⁷ at 42 °C. For *secY24*-bearing *E. coli* strains²¹, suppression tests were performed with the *groES*⁺ *EL*⁺-bearing plasmid pOF39 (ref. 7) on minimal M9 medium¹⁷ supplemented with glucose and ampicillin. Each *S. typhimurium* and *E. coli* strain contained a wild-type chromosomal copy of the *groE* operon. Abbreviations: CS, Cold-sensitive; HS, heat-sensitive; TI, temperature-independent; NA, not applicable.

* The *his* alleles, numbered below 1,000 are reviewed in ref. 13 whereas those numbered above 1,000 are described in ref. 14.

† The *ilv* alleles have not been reported previously.

‡ Ref. 23.

§ Ref. 22.

|| Ref. 20.

¶ Ref. 21.

Ref. 18.

of the activity present in the wild-type strain containing the plasmid overexpressing *groESL*.

All but one of the enzymes involved in histidine biosynthesis are multimeric¹⁵. Only two heat-sensitive alleles of *hisA* (ref. 14), specifying the monomeric protein phosphoribosylformimino-5-amino-1-phosphoribosyl-4-imidazolecarboxamide isomerase¹⁵ exist; neither was suppressed by the multicopy *groE*-bearing plasmid (Table 1). Heat-sensitive mutants caused by alteration of monomeric proteins, although rare, have been described for several aminoacyl-tRNA synthetases in *E. coli*¹⁸. The proteins of this family display diverse quaternary structures (α , α_2 , $\alpha\beta$, $\alpha_2\beta_2$, α_4)¹⁹. The *groE*-bearing plasmid did not suppress heat-sensitive growth defects caused by mutations in

alaS, *glnS*, *leuS*, *metG* and *serS*, which are structural genes encoding, respectively, alanyl-(α_4), glutamyl-(α), leucyl-(α), methionyl-(α_2) and seryl-(α_2) tRNA synthetases in *E. coli* (Table 1).

Protein translocation involves interactions among several components. Among mutations affecting the mitochondrial protein import machinery in *Saccharomyces cerevisiae* are alleles of *MIF4*, which encodes a protein homologue of *E. coli* GroEL⁴. We tested the ability of the *groE*-bearing plasmid to suppress heat-sensitive growth (Fig. 1; Table 1) and secretion (A.A.G., unpublished data) defects caused by the *secA51*²⁰ and *secY24*²¹ alleles of *E. coli*. Suppression of both alleles was observed (Table 1), indicating that the suppression mechanism is gene-independent in *E. coli* as well as in *S. typhimurium*. The presence of pGroESL had no effect on the growth of a strain containing a conditionally lethal amber mutation in gene *X* upstream of *secA* (A.A.G., unpublished data); the lethality is due to the polar effect of the mutation on *secA*, resulting in the absence of the *secA* protein²⁰. This implies that the *secA* protein must be present for the suppression of the secretion defect and that *groE* overexpression cannot bypass this requirement.

Can mutations altering structural proteins, rather than enzymes, be suppressed by overexpression of the *groE* operon products. Gene 9 of *Salmonella* phage P22 encodes a polypeptide that trimerizes to form the phage's tail spike, six of which are attached to each phage particle²². Many extensively characterized heat-sensitive protein-folding mutations of gene 9 exist^{22,23}. Phages containing such mutations do not plaque on wild-type *S. typhimurium* at the restrictive temperature. We determined the ability of these mutants to form plaques on strains containing either a control or the *groE* multicopy plasmid at the restrictive temperature (Table 1). Several mutations were suppressed.

We characterized the suppression of the P22 *9tsfH304* mutation more fully. The efficiency of plating of this mutant at 39 °C versus 30 °C on a strain containing a control plasmid³ was less than 10⁻⁴. By contrast, the efficiency of plating on a host containing the plasmid overexpressing *groESL* (ref. 3) was 0.1. At 30 °C we observed a burst size of ~200 with either host. Infection of the strain overexpressing *groESL* with the mutant phage resulted in a burst of six plaque-forming units (PFU) at 39 °C; a burst of 0.4 PFU was observed when the host contained the control plasmid.

The nucleotide and corresponding deduced amino-acid sequences of the *groE*-suppressed phage P22 gene 9 (refs 22,

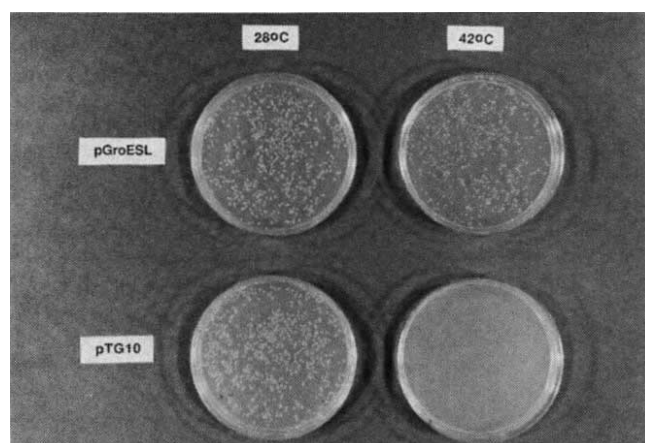


FIG. 1 Suppression of the heat-sensitive phenotype of a *secA51* strain. Strain MM52 (*secA51*)²⁰ was transformed with either plasmid pTG10 (*groES*⁻ *groEL*⁻) or plasmid pGroESL (*groES*⁺ *groEL*⁺) selecting for chloramphenicol resistance. Recombinants were plated for single colonies on antibiotic-containing LB medium at the indicated temperatures. The presence of plasmid pGroESL suppressed the heat-sensitive phenotype of the *secA51* mutant.

TABLE 2 Nucleotide sequences and deduced amino-acid sequences of suppressible alleles

Allele	Codon change	Amino-acid change	Amino-acid residue	Ref.
<i>secA51</i>	CTG → CCG	Leu → Pro	43	20
<i>secY24</i>	GGT → GAT	Gly → Asp	240	21
P22 9 <i>tsfU34</i>	AAG → GAG	Lys → Glu	163	23
P22 9 <i>tsfH304</i>	GGA → AGA	Gly → Arg	244	22
P22 9 <i>tsfU3</i>	GAA → GTA	Glu → Val	309	23
P22 9 <i>tsfU8</i>	GAA → AAA	Glu → Lys	344	23

The *E. coli secA*, *secY* and the phage P22 gene 9 proteins are of 901, 443 and 666 amino-acid residues, respectively²⁰⁻²³.

23) and the *sec* alleles^{20,21} are known (Table 2). At the protein level, these gene 9 transition and transversion mutations cause replacement of an acidic side chain by a basic one (glutamic acid → lysine) and a basic R-group by an acidic one (lysine → glutamic acid), loss of a negative charge (glutamic acid → valine) and substitution of a bulky positively charged side chain for a proton (glycine → arginine). The *secA* mutation substitutes the imino acid proline in place of leucine, whereas the *secY* allele results in the substitution of the acidic side chain of aspartic acid for the proton of glycine. Proposed suppression mechanisms must be consistent with the wide variety of suppressible alleles.

Suppression of *dnaA* alleles^{7,8}, phage λ morphogenesis¹⁰, and ribulose biphosphate carboxylase assembly³, require both *groEL* and *groES*. Both *groE* genes were also needed to mediate suppression of the mutations used in our study. Plasmids expressing either *groEL* or *groES* were unable to suppress representative alleles of gene 9, *hisD*, *hisC*, *ilvGM* and *secA* responsive to the intact *groESL* operon (Table 3).

There are several mechanisms by which multiple copies of *groE* could act as a general suppressor of heat-sensitive and perhaps other missense mutations. *GroE* operon products could alter the translational fidelity by their interaction with the ribosome²⁴, or reduce protein degradation by having an anti-proteolytic role, or modulate transcription rate, translation rate or ionic strength¹⁴. But *groE*-mediated translational misreading is unlikely to be a common means of suppression, because misreading, induced by the presence of aminoglycoside antibiotics, does not allow growth of the *groE*-suppressible *hisC15* mutant²⁵. Also, the broad spectrum of suppressible alleles (Table 2) precludes codon-specific mechanisms of informational suppression. And because protein degradation is decreased in a *groEL* mutant of *E. coli*²⁶, it is unlikely that *GroE* over-production antagonizes degradation. Elevated transcription or translation of target genes also cannot explain the observed suppression, because expression of the *Salmonella his* and *ilv* operons is not increased by the presence of the multicopy *groE*-bearing plasmid pGroESL (R.A.L., unpublished observations). Finally, the *groE*-suppressible *ilvGM3139* allele is not responsive to phenotypic suppression by sodium chloride

TABLE 3 Suppression requires *groES* and *groEL*

Allele	pTG10 (<i>groES</i> ⁻ L ⁻)	Growth in the presence of: pGroEL (<i>groES</i> ⁻ L ⁺)	pGroES (<i>groES</i> ⁻ L ⁻)	pGroESL (<i>groES</i> ⁺ L ⁺)
<i>hisC6595</i>	—	—	—	+
<i>hisD6592</i>	—	—	—	+
<i>ilvGM3139</i>	—	—	—	+
<i>secA51</i>	—	—	—	+
P22 9 <i>tsfH304</i>	—	—	—	+

Plasmid pGroES was constructed by ligating the small *EcoRI* restriction fragment from plasmid pS4 (ref. 7) into the *EcoRI* site of the polylinker present in pTG10 (ref. 3). Plasmid pGroES and the other listed plasmids³ were introduced into a variety of strains carrying the indicated chromosomal mutation and the host for phage P22 growth, DB7136, as described in the legend to Table 1. Conditions to test suppression were as described in Table 1.

supplementation¹⁴ (T.K.V.D., unpublished observations), indicating that elevation of the internal ionic strength cannot be a common means of suppressing the *groE*-responsive heat-sensitive mutations. Thus, the explanation most consistent with present views^{1,27} and our data is that overexpressed *groE* gene products directly interact with the mutant proteins, allowing them to achieve their active conformations. Such interactions could reflect a normal association of the *GroE* proteins with the translation products of many genes, including wild-type versions of the suppressed alleles. □

Received 23 June; accepted 2 October 1989.

- Hemmingsen S. M. *et al.* *Nature* **333**, 330–334 (1988).
- Friedman D. I. *et al.* *Microbiol. Rev.* **48**, 299–325 (1984).
- Goluboinoff P., Gatenby, A. A. & Lorimer, G. H. *Nature* **337**, 44–47 (1989).
- Cheng, M. Y. *et al.* *Nature* **337**, 620–625 (1989).
- Fayet, O., Ziegelhoffer, T. & Georgopoulos, C. *J. Bact.* **171**, 1379–1385 (1989).
- Neidhardt, F. C. & van Bogelen, R. A. in *Escherichia coli and Salmonella typhimurium* (eds Neidhardt, F. C. *et al.*) 1334–1345 (American Society for Microbiology, Washington, D.C., 1987).
- Fayet, O., Louarn, J.-M. & Georgopoulos, C. *Molec. Gen. Genet.* **202**, 435–445 (1986).
- Jenkins, A. J., March, J. B., Oliver, I. R. & Masters, M. *Molec. Gen. Genet.* **202**, 446–454 (1986).
- McMacken, R., Silver, L. & Georgopoulos, C. in *Escherichia coli and Salmonella typhimurium* (eds Neidhardt, F. C. *et al.*) 564–612 (American Society for Microbiology, Washington, D.C., 1987).
- Tilly, K., Murialdo, H. & Georgopoulos, C. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1629–1633 (1981).
- Srivastava, D. K. & Bernhard, S. A. *Science* **234**, 1081–1086 (1986).
- Hong, J.-S. & Ames, B. N. *Proc. natn. Acad. Sci. U.S.A.* **68**, 3158–3162 (1971).
- Hartman, P. E., Hartman, Z., Stahl, R. C. & Ames, B. N. *Adv. Genet.* **16**, 1–34 (1971).
- Kohn, T. & Roth, J. *Biochemistry* **18**, 1386–1392 (1979).
- Winkler, M. E. in *Escherichia coli and Salmonella typhimurium* (eds Neidhardt, F. C. *et al.*) 395–411 (American Society for Microbiology, Washington, D.C., 1987).
- Ciesla, Z., Salvatore, F., Broach, J. R., Artz, S. W. & Ames, B. N. *Analyt. Biochem.* **63**, 44–55 (1975).
- Davis, R. W., Botstein, D. & Roth J. *Advanced Bacterial Genetics* (Cold Spring Harbor Laboratory, New York 1980).
- LaRossa, R. & Soll, D. in *Transfer RNA* (ed. Altman, S.) 136–167 (MIT, Cambridge, Massachusetts, 1978).
- Schimmel, P. A. *Rev. Biochem.* **56**, 125–158 (1987).
- Schmidt, M. G., Rollo, E. E., Grodberg, J. & Oliver, D. B. *J. Bact.* **170**, 3404–3414 (1988).
- Shiba, K., Ito, K., Yura, T. & Cerretti, D. P. *EMBO J.* **3**, 631–635 (1984).
- Yu, M.-H. and King, J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6584–6588 (1984).
- Villafane, R. & King, J. *J. molec. Biol.* **204**, 607–619 (1988).
- Neidhardt, F. C. *et al.* *A. R. J. Bact.* **145**, 513–520 (1981).
- Whitfield, Jr., H. J., Martin, R. G. & Ames, B. N. *J. molec. Biol.* **21**, 335–355 (1966).
- Straus, D. B., Walter, W. A. & Gross, C. A. *Genes Dev.* **2**, 1851–1858 (1988).
- Bochkareva, E. S., Lissin, N. M. & Girshovich, A. S. *Nature* **336**, 254–257 (1988).
- Enomoto, M. & Stocker, B. A. D. *Virology* **60**, 503–514 (1974).
- Sanderson, K. E. & Roth, J. R. *Microbiol. Rev.* **47**, 410–453 (1983).
- Fane, B. & King, J. *Genetics* **117**, 167–171 (1987).

ACKNOWLEDGEMENTS. We thank Philip Hartman, Jonathan King, Cameron Haase-Pettingel and John Roth for provision of strains, exchange of ideas and encouragement, and Anca Segall, Nat Sternberg and Paul Viitanen for useful comments. The *sec* mutants used in this study were gifts from Koreaki Ito and Donald Oliver.

Identification of a novel translation factor necessary for the incorporation of selenocysteine into protein

Karl Forchhammer, Walfred Leinfelder & August Böck

Lehrstuhl für Mikrobiologie der Universität, Maria-Ward-Strasse 1a, D-8000 München 19, FRG

DURING the biosynthesis of selenoproteins in both prokaryotes and eukaryotes, selenocysteine is cotranslationally incorporated into the nascent polypeptide chain^{1,2} through a process directed by a UGA codon that normally functions as a stop codon³⁻⁵. Recently, four genes have been identified whose products are required for selenocysteine incorporation in *Escherichia coli*⁶. One of these genes, *selC*, codes for a novel transfer RNA species (tRNA^{UCA}) that accepts serine and cotranslationally inserts selenocysteine by recognizing the specific UGA codon⁷. The serine residue attached to this tRNA is converted to selenocysteine in a reaction dependent on functional *selA* and *selD* gene products⁸. By contrast, the *selB* gene product (SELB) is not required until after selenocysteyl-tRNA biosynthesis⁸. Here we present evidence indicating that SELB is a novel translation factor. The deduced

amino-acid sequence of SELB exhibits extensive homology with the sequences of the translation initiation factor-2 (IF-2) and elongation factor Tu (EF-Tu). Furthermore, purified SELB protein binds guanine nucleotides in a 1:1 molar ratio and specifically complexes selenocysteyl-tRNA^{UCA}, but does not interact with seryl-tRNA^{UCA}. Thus, SELB could be an amino acid-specific elongation factor, replacing EF-Tu in a special translational step.

Functional SELB protein is absolutely required for translation of the selenocysteine-specific UGA codon in the *fdhF* messenger RNA (ref. 6, and unpublished results). We first determined the nucleotide sequence of the *selB* gene (Fig. 1), and found that it potentially encodes a protein of relative molecular mass 68,800 (*M_r*, 68.8K) consisting of 614 amino acids. Computer-assisted analysis of the amino-acid sequence revealed a striking similarity of the N-terminal 244 amino acids to the sequence of the two translation factors EF-Tu and IF-2 (Fig. 2). It is of interest that the homologous region between SELB and EF-Tu corresponds to the segment of primary structure that also is conserved between EF-Tu and IF-2. In contrast to all G proteins identified so far, SELB shows one unique deviation in its guanine-nucleotide-binding motif⁹—the universal asparagine residue (position 135 in EF-Tu), which forms a hydrogen bond with the keto group of the guanine ring¹⁰, is replaced by threonine. Sequence similarity was also found between SELB and several other G proteins, particularly EF-G, LEPA and p21. In these cases, however, similarity is restricted to the first 120 amino acids in SELB, corresponding to the G domain of EF-Tu (data not shown).

EF-Tu and IF-2 share the property of binding guanine nucleotides and aminoacyl-tRNA. If it is assumed that this function is associated with the conserved domain, then it might be anticipated that SELB exhibits similar properties. To examine this possibility, the *selB* gene product was overproduced in a T7 promoter expression system and purified (K.F., K. P. Rücknagel and A.B., manuscript in preparation). We then used equilibrium dialysis for nucleotide-binding studies with purified SELB (Fig. 3). The SELB protein bound guanine nucleotides in a 1:1 molar ratio. The dissociation constants were calculated to be 10 μ M for SELB-GDP and 1.7 μ M for SELB-GTP (Fig. 3a, b). Binding is specific for guanine nucleotides, because the addition of a 10-fold excess of various other nucleotides did not result in competition for binding of [³H]GTP (Fig. 3c); binding of [³H]GDP could only be quenched by GTP (data not shown). Thus, SELB binds GDP with an affinity similar to that exhibited by IF-2 (ref. 11). This deviates significantly from the affinity of EF-Tu for GDP, which is three orders of magnitude greater¹². The trend is reversed for GTP binding; SELB and EF-Tu have similar affinities for GTP, whereas the affinity of IF-2 for GTP is lower by a factor of ~100 (refs 11, 12).

To examine whether SELB also interacts with *selC*-encoded tRNA^{UCA}, we purified the latter from a strain carrying the gene on a multicopy plasmid²³. We analysed complex formation by determining the protection of the tRNA by SELB either against alkaline hydrolysis of the aminoacyl-tRNA ester bond^{13,14}, or against enzymatic cleavage of the tRNA by RNase¹⁵. After ligation to tRNA^{UCA}, L-serine was converted *in vitro* to selenocysteyl-tRNA by using extracts from cells overproducing the *selA* and *selD* gene products²³. We then attempted to chemically deacylate the recovered selenocysteyl-tRNA in the presence or absence of SELB. We found a significant degree of protection of the tRNA against hydrolysis in the presence of SELB. This effect was independent of the addition of guanine nucleotides (Fig. 4a). By contrast, we could not detect any protection from deacylation when seryl-tRNA^{UCA} was used in the same assay (Fig. 4b). We also found no protection with any other aminoacyl-tRNA tested (data not shown). Similar results were obtained when we assayed complex formation using the RNase-protection method¹⁵. SELB specifically protected selenocysteyl-tRNA^{UCA} (Fig. 4c, d). Again, this effect did not require the presence of guanine nucleotides.

-50	GATTTACGGT GCCTTGAAGA TGACCAACGG TTTTGGAGA TGTTGTTGAA																			
1	ATG	ATT	ATT	GCG	ACT	GCC	GGA	CAC	GTT	GAC	CAC	GGC	AAA	ACA	ACC	TTA	TTG	CAG		
1	MET	Ile	Ile	Ala	Thr	Ala	Gly	His	Val	Asp	His	Gly	Lys	Thr	Thr	Leu	Leu	Gln		
55	GCG	ATT	ACT	GCG	GTA	AAT	GCT	GAC	CGT	CTG	CCG	GAA	GAA	AAA	AAG	CGC	GCC	ATG		
19	Ala	Ile	Thr	Gly	Val	Asn	Ala	Asp	Arg	Leu	Pro	Glu	Glu	Lys	Lys	Arg	Gly	MET		
109	ACC	ATC	GAT	CTC	GCG	TAT	GCC	TAC	TGG	CCG	CAG	CCG	GAT	GGT	CGC	GTG	CCT	GGT		
37	Thr	Ile	Asp	Leu	Gly	Tyr	Ala	Tyr	Trp	Pro	Gln	Pro	Asp	Gly	Arg	Val	Pro	Gly		
163	TTT	ATC	GAC	GTT	CCC	GGT	CAT	GAA	AAG	TTT	CTT	TCC	AAC	ATG	CTG	GCG	GCC	GTT		
55	Phe	Ile	Asp	Val	Pro	Gly	His	Glu	Lys	Phe	Leu	Ser	Asn	MET	Leu	Ala	Gly	Val		
217	GGT	GGT	ATC	GAT	CAC	GCG	CTG	TTG	GTG	GTG	GCG	TGC	GAT	GAC	GCG	GTG	ATG	GCA		
73	Gly	Gly	Ile	Asp	His	Ala	Leu	Leu	Val	Val	Ala	Cys	Asp	Asp	Gly	Val	MET	Ala		
271	CAG	ACC	CGT	GAG	CAT	CTG	GCG	ATT	TTG	CAG	CTG	ACC	GGT	AAC	CCG	ATG	CTG	ACA		
91	Gln	Thr	Arg	Glu	His	Leu	Ala	Ile	Leu	Gln	Leu	Thr	Gly	Asn	Pro	MET	Leu	Thr		
325	GTG	GCG	CTG	ACC	AAA	GCC	GAT	CGC	GTG	GAC	GAA	GCG	CGT	GTT	GAT	GAG	GTT	GAA		
109	Val	Ala	Thr	Lys	Lys	Ala	Asp	Arg	Val	Asp	Glu	Ala	Arg	Val	Asp	Glu	Val	Glu		
379	CGC	CAG	GTA	AAG	GAG	GTT	CTG	GCG	GAA	TAC	GGT	TTT	GCT	GAG	GCA	AAA	CTG	TTT		
127	Arg	Gln	Val	Lys	Glu	Val	Leu	Arg	Glu	Tyr	Gly	Phe	Ala	Glu	Ala	Lys	Leu	Phe		
433	ATC	ACC	GCA	ACC	GAA	GCG	GAT	GCG	GGA	ATG	GAT	GCC	CTG	GCG	GAG	CAT	CTG	CTT		
145	Ile	Thr	Ala	Ala	Thr	Glu	Gly	Arg	Gly	MET	Asp	Ala	Leu	Arg	Glu	His	Leu	Leu		
487	CAG	TTG	CCG	GAA	CGC	GAG	CAC	GCC	AGC	CAA	CAT	AGT	TTC	CGC	CTC	GCG	ATT	GAC		
163	Gln	Leu	Pro	Glu	Arg	Glu	His	Ala	Ser	Gln	His	Ser	Phe	Arg	Leu	Ala	Ile	Asp		
541	CGC	GCA	TTT	ACC	GTA	AAA	GGT	GCC	GCG	CTG	GTC	GTC	ACC	GGT	ACG	GCG	TTA	AGC		
181	Arg	Ala	Phe	Thr	Val	Lys	Gly	Ala	Gly	Leu	Val	Val	Thr	Gly	Thr	Ala	Leu	Ser		
595	GGG	GAA	GTG	AAG	GTA	GCG	GAT	TCA	CTC	TGG	CTG	ACT	GGT	GTA	AAT	AAA	CCG	ATG		
199	Gly	Glu	Val	Lys	Val	Gly	Asp	Ser	Leu	Trp	Leu	Thr	Gly	Val	Asn	Lys	Pro	MET		
649	CGT	GTA	CGT	GCG	CTG	CAT	GCG	CAA	AAC	CAG	CCA	ACA	GAA	ACC	GCC	AAT	GCC	GGG		
217	Arg	Val	Arg	Ala	Leu	His	Ala	Gln	Ala	Asn	Gln	Pro	Thr	Glu	Thr	Ala	Ile	Gly		
703	CAG	CGT	ATC	GCG	CTT	AAC	ATC	GCG	GGT	GAT	GCG	GAA	AAA	GAG	CAG	ATT	AAC	CGT		
235	Gln	Arg	Ile	Ala	Leu	Asn	Ile	Ala	Gly	Asp	Ala	Glu	Lys	Glu	Gln	Ile	Asn	Arg		
757	GCG	GAC	TGG	CTG	CTT	GCC	GAT	GTG	CCG	GCA	GAG	CGC	TTC	ACA	GCG	GTG	ATT	GTC		
253	Gly	Asp	Trp	Leu	Leu	Ala	Asp	Val	Pro	Pro	Glu	Pro	Phe	Thr	Arg	Val	Ile	Val		
811	GAG	CTT	CAA	ACC	CAT	ACA	CCG	CTG	ACC	CAG	TGG	CAG	CCG	CTG	CAT	ATT	CAC	CAC		
271	Glu	Leu	Gln	Thr	His	Thr	Pro	Leu	Thr	Gln	Trp	Gln	Pro	Leu	His	Ile	His	His		
865	GCC	GCC	AGC	CAC	GTC	ACG	GGA	CGC	GTT	TCA	CTG	CTG	GAA	GAT	AAC	CTT	GCT	GAA		
289	Ala	Ala	Ser	His	Val	Thr	Gly	Arg	Val	Ser	Leu	Glu	Asp	Asn	Leu	Ala	Glu			
919	CTG	GTC	TTC	CAC	ACC	CCG	TAA	TGG	CTG	GCA	GAT	AAC	CAC	CGC	CTG	GTA	TTG	GCG		
307	Leu	Val	Phe	Asp	Thr	Pro	Leu	Trp	Leu	Ala	Asp	Asn	Asp	Arg	Leu	Val	Leu	Arg		
973	GAT	ATC	TCT	GCC	CGC	AAC	ACG	CTG	GCC	GGA	GCG	CGC	GTC	GTG	ATG	CTT	AAC	CCG		
325	Asp	Ile	Ser	Ala	Arg	Asn	Thr	Leu	Ala	Gly	Val	Ala	Met	Leu	Asn	Pro				
1027	CCG	CGT	CGC	GGT	AAA	CGT	AAG	CCG	GAA	TAT	CTG	CAA	TGG	CTG	GCG	TCT	CTT	GCA		
343	Pro	Arg	Arg	Gly	Lys	Arg	Lys	Pro	Glu	Tyr	Leu	Gln	Trp	Leu	Gln	Thr	Ala	Leu		
1081	CGG	GCG	GAG	AGC	GAT	GCC	GAT	GCG	TAT	TCT	GTT	CTG	Pro	GAA	CGC	GCG	GCG	GTT		
361	Arg	Ala	Gln	Ser	Asp	Ala	Asp	Ala	Leu	Ser	Val	His	Leu	Glu	Arg	Gly	Ala	Val		
1135	AAC	CTT	GCG	GAT	TTC	GCC	TGG	GCG	CGC	CAG	CTC	AAC	GCG	GAA	GGG	ATG	CGC	GAA		
379	Asn	Leu	Ala	Asp	Phe	Ala	Trp	Ala	Arg	Gln	Leu	Asn	Gly	Glu	Gly	MET	Arg	Glu		
1189	TTG	CTG	CAA	CAG	CCT	GGT	TAT	ATT	CAG	GCT	GGT	TAT	AGC	TTG	TTG	AAT	GCG	CCG		
397	Leu	Leu	Gln	Gln	Pro	Gly	Tyr	Ile	Gln	Ala	Gly	Tyr	Ser	Leu	Leu	Asn	Ala	Pro		
1243	GTT	GCC	GCC	CGC	TGG	CAG	CGG	AAA	ATT	CTC	GAC	ACA	TTA	GCG	ACT	TAT	CAT	GAG		
415	Val	Ala	Ala	Arg	Trp	Gln	Arg	Lys	Ile	Leu	Asp	Thr	Leu	Ala	Thr	Tyr	His	Glu		
1297	CAA	CAT	CGC	GAT	GAA	CCT	GGC	CCT	GGG	CGC	GAA	CGT	CTG	GCA	CGT	ATG	GCG	TTG		
433	Gln	His	Arg	Asp	Glu	Pro	Gly	Pro	Gly	Arg	Leu	Arg	Arg	Met	Ala	Leu				
1351	CCA	ATG	GAA	GAT	GAA	GCG	CTG	GTA	CTG	TTG	CTG	ATT	GAA	AAG	ATG	CGC	GAA	AGC		
451	Pro	MET	Glu	Asp	Glu	Ala	Leu	Val	Leu	Leu	Leu	Ile	Glu	Lys	MET	Arg	Glu	Ser		
1405	GCG	GAC	ATC	CAC	AGC	CAT	CAC	GCG	TGG	CTG	CAT	CTG	CCA	GAT	CAC	AAA	GCG	GCG		
469	Gly	Asp	Ile	His	Ser	His	His	Gly	Trp	Leu	His	Leu	Pro	Asp	His	Lys	Ala	Gly		
1459	TTC	AGC	GAA	GAG	CAG	CAG	GCC	ATC	TGG	CAA	AAA	GCA	GAG	CCA	CTG	TTT	GGT	GAC		
487	Phe	Ser	Glu	Glu	Gln	Gln	Ala	Ile	Trp	Gln	Lys	Ala	Glu	Pro	Leu	Phe	Gly	Asp		
1513	GAA	CCG	TGG	TGG	CTG	CGT	GAC	CTG	GCA	AAA	GAG	ACG	GGA	ACC	GAC	GAG	CAG	GCA		
505	Glu	Pro	Trp	Trp	Val	Arg	Asp	Leu	Ala	Lys	Glu	Thr	Gly	Thr	Asp	Glu	Gln	Ala		
1567	ATG	CGC	CTG	ACT	CTA	GCG	CAG	GCG	GCG	CAG	CAA	GGA	ATA	ATT	ACC	GCG	ATC	GTT		
523	MET	Arg	Leu	Thr	Leu	Arg	Gln	Ala	Ala	Gln	Gln	Gly	Ile	Ile	Thr	Ala	Ile	Val		
1621	AAA	GAT	CGT	TAT	TAC	CGT	AAC	GAT	CGG	ATT	GTC	GAG	TTT	GCC	AAT	ATG	ATC	CGC		
541	Lys	Asp	Arg	Tyr	Tyr	Arg	Asn	Asp	Arg	Ile	Val	Glu	Phe	Ala	Asn	MET	Ile	Arg		
1675	GAT	CTC	GAT	CAG	GAG	TGT	GGT	TCA	ACC	TGC	GCG	GAT	TTT	CGC	GAT	CGC	TTA			
559	Asp	Leu	Asp	Gln	Glu	Cys	Gly	Ser	Thr	Cys	Ala	Ala	Asp	Phe	Arg	Asp	Arg	Leu		
1729	GGC	GTA	GCG	GCA	AAG	CTG	GCA	ATT	CAG	ATT	CTG	GAA	TAT	TTT	GAC	CGC	ATT	GCG		
577	Gly	Val	Gly	Arg	Lys	Leu	Ala	Ile	Gln	Ile	Gln	Leu	Tyr	Phe	Ala	Leu	Ile	Gly		
1783	TTT	ACG	CGT	CGT	CGT	GGA	AAT	GAT	CAT	TTA	TTG	CGC	GAC	GCA	TTA	TTA	TTT	CCG		
595	Phe	Thr	Arg	Arg	Arg	Gly	Asn	Asp	His	Leu	Leu	Arg	Asp	Ala	Leu	Leu	Phe	Pro		
1837	GAA	AAA	TAA	GGAAATGATT																
613	Glu	Lys	***																	

FIG. 1 Nucleotide sequence of the *selB* gene and its derived amino-acid sequence. The 2.1 kilobase (kb) insert of plasmid pWL144 (ref. 6) containing the *selB* gene was sequenced by the method of Maxam and Gilbert¹⁶ using either restriction sites within the plasmid or the endpoints of nested Bal31 deletion clones. The putative Shine-Dalgarno sequence is underlined and the termination codon is indicated by asterisks. The sequence data are deposited in the EMBL database under the accession number X16644.

These results strongly indicate that SELB specifically complexes selenocysteyl-tRNA^{UCA} but does not recognize tRNA^{UCA} aminoacylated with serine. Complex formation is independent of the presence of guanine nucleotides, as is complex formation with *E. coli* IF-2 (ref. 14) or EF-Tu from thermophilic bacteria (M. Sprinzl, personal communication). The basis for this specificity towards selenocysteine remains to be elucidated, but could depend on either the tRNA^{UCA} undergoing a conformational change when serine is converted to selenocysteine, or, more plausibly, the direct recognition by SELB of the selenol group at the aminoacyl residue.

SELB most probably functions biochemically as an alternative elongation factor to EF-Tu, specific for the insertion of selenocysteine directed by the UGA codon in the *fdhF* mRNA. Apart from the sequence homology with EF-Tu and IF-2, several other lines of evidence support this role of SELB. First, SELB binds guanine nucleotides and selenocysteyl-tRNA. Second, EF-Tu exhibits a 200-fold lower affinity for seryl-tRNA^{UCA} than for any other aminoacyl-tRNA involved in elongation (G. Ott and M. Sprinzl, personal communication). Furthermore, conversion of the serine residue to a selenocysteine residue does not change this affinity (G. Ott, K.F. and M. Sprinzl, unpublished results). Thus, under *in vivo* conditions, selenocysteyl-tRNA cannot compete for ternary complex formation with EF-Tu. Finally, lesions in *selB* do not interfere with selenocysteyl-tRNA formation⁸; rather, they promote an accumulation of selenocysteyl-tRNA and, despite the resulting intracellular selenocysteyl-tRNA concentration, there is no incorporation of selenocysteine at the specific UGA codon by EF-Tu.

These findings are therefore consistent with SELB being a new G protein functioning as a translation factor that specifically recognizes the aminoacyl residue of a tRNA. The recognition of the selenocysteyl residue esterified to tRNA^{UCA} thus could be the biochemical basis for the extreme specificity with which the UGA codon in the *fdhF* mRNA directs the insertion of selenocysteine only, and not the insertion of L-serine; this is exemplified in mutants that cannot convert seryl-tRNA^{UCA} into selenocysteyl-tRNA⁸. The higher *M_r* of SELB (68K) relative to that of EF-Tu (43K) indicates that SELB probably has additional functions. One function of SELB could be to recognize the aminoacyl residue. Other functions could be the specific recognition of the mRNA context around the selenocysteine-specific UGA codon and/or to compete for the binding of release factor-2. □

The conformation of the DNA double helix in the crystal is dependent on its environment

Zippora Shakked, Gali Guerstein-Guzikevich,
Miriam Eisenstein, Felix Frolov & Dov Rabinovich

Department of Structural Chemistry, The Weizmann Institute of Science, Rehovot 76100, Israel

STUDIES of the crystal structures of more than 30 synthetic DNA fragments have provided structural information about three basic forms of the double helix: A-, B- and Z-form DNA¹⁻⁵. These studies have demonstrated that the DNA double helix adopts a highly variable structure which is related to its base sequence. The extent to which such observed structures are influenced by the crystalline environment can be found by studying the same molecule in different crystalline forms. We have recently crystallized one particular oligomer in various crystal forms. Here we report the results of structural analyses of the different crystal structures and demonstrate that the DNA double helix can adopt a range of conformations in the crystalline state depending on hydration, molecular packing and temperature. These results have implications on our understanding of the influence of the environment on DNA structure, and on the modes of DNA recognition by proteins.

The octamer d(G-G-G-C-G-C-C-C) or G-G-G-C-G-C-C-C has been crystallized as an A-DNA double helix in two crystal forms, tetragonal and hexagonal. The structure of each form was determined at room temperature (RT) and low temperature (LT). The experimental details for crystallization, data collection and structure determination are given in Table 1.

The crystal structures of the tetragonal forms have been previously reported^{6,7} and their double-helical structures at the two temperatures found to be very similar; therefore, we refer to only the RT structure of the tetragonal form. We found that when hexagonal crystals are cooled, however, there are drastic changes in unit-cell volume and resolution of the data (Table 1). Such changes indicate a phenomenon of phase transition, and therefore the crystalline environment of the hexagonal LT structure is different from that of the RT structure. Some of the global and local parameters of each of the three structures of the same sequence are given in Table 2, and continuous helices based on the three structures displayed in Fig. 1. Each of these three crystal structures corresponds to a distinct conformation of the A-DNA double helix. The tetragonal structure is of conformation I. This is characterized by a very small tilting (inclination) of the base pairs with respect to the overall helix axis, a large rise per base-pair and a widely open major groove. The hexagonal RT structure is of conformation II, with medium values of base tilting and rise. The major groove is narrower and deeper than that of the tetragonal structure. The hexagonal LT structure is of conformation III. This is a slim helix with a narrower and shallower major groove than those of conformations I and II. The tilting of the bases and the rise are similar to those of conformation II.

These data show that the same oligomer can assume more than one conformation in the crystalline state. The three helices are embedded in three distinct crystalline environments. Among the factors that are likely to influence molecular conformation and crystal packing are hydration and temperature. Analysis of the water structure in the three crystals shows that the level and mode of hydration changes considerably from one conformational type to the next. Conformation I is highly hydrated, II is poorly hydrated and III is fairly hydrated. It seems that during the cooling process of the hexagonal crystals their water structure undergoes an important rearrangement which affects both the helical conformation and packing arrangement.

Received 24 July; accepted 2 October 1989.

1. Zinoni, F., Birkmann, A., Leinfelder, W. & Böck, A. *Proc. natn. Acad. Sci. U.S.A.* **84**, 3156-3160 (1987).
2. Böck, A. & Stadtmann, T. C. *Biofactors* **1**, 245-250 (1988).
3. Zinoni, F., Birkmann, A., Stadtmann, T. C. & Böck, A. *Proc. natn. Acad. Sci. U.S.A.* **83**, 4650-4654 (1986).
4. Chambers, I. et al. *EMBO J.* **5**, 1221-1227 (1986).
5. Suenaga, Y., Ishida, K., Takeda, T. & Takagi, K. *Nucleic Acids Res.* **15**, 7187 (1987).
6. Leinfelder, W. et al. *J. Bacteriol.* **170**, 540-546 (1988).
7. Leinfelder, W., Zehlein, E., Mandrand-Berthelot, M. A. & Böck, A. *Nature* **331**, 723-725 (1988).
8. Leinfelder, W., Stadtmann, T. C. & Böck, A. *J. Biol. Chem.* **264**, 9720-9723 (1989).
9. Dever, T. E., Glynnias, M. J. & Merrick, W. C. *Proc. natn. Acad. Sci. U.S.A.* **84**, 1814-1818 (1987).
10. LaCour, T. M. F., Nyborg, J., Thirpu, S. & Clark, B. F. C. *EMBO J.* **4**, 2385-2388 (1985).
11. Gualerzi, C. O. et al. in *Structure, Function, and Genetics of Ribosomes* (eds Hardesty, B. & Kramer, G.) 621-641 (Springer, New York, 1986).
12. Arai, K., Kawakita, M. & Kaziro, Y. *J. Biochem., Tokyo* **76**, 293-306 (1976).
13. Pingoud, A., Urbanke, C., Krauss, G., Peters, F. & Maass, G. *Eur. J. Biochem.* **78**, 403-409 (1977).
14. Petersen, H. U., Roll, T., Grunberg-Manago, M. & Clark, B. F. C. *Biochem. biophys. Res. Commun.* **91**, 1068-1074 (1979).
15. Knowlton, R. G. & Yarus, M. *J. Molec. Biol.* **139**, 721-732 (1980).
16. Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499-561 (1980).
17. Jones, M. D. et al. *Eur. J. Biochem.* **108**, 507-526 (1980).
18. Sacerdot, C. P., Dessen, P., Hershey, J. W. B., Plumbbridge, J. A. & Grunberg-Manago, M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7787-7791 (1984).
19. Kohno, K. et al. *Proc. natn. Acad. Sci. U.S.A.* **83**, 4978-4982 (1986).
20. Van Noort, J. M. et al. *Eur. J. Biochem.* **160**, 559-561 (1986).
21. Jonak, J., Pokorna, K., Meloun, B. & Karas, K. *J. Biochem.* **154**, 355-362 (1986).
22. Van Noort, J. M., Kraal, B. & Bosch, L. *Proc. natn. Acad. Sci. U.S.A.* **82**, 3212-3216 (1985).
23. Leinfelder, W., Forchhammer, K., Veprek, B., Zehlein, E. & Böck, A. *Proc. natn. Acad. Sci. U.S.A.* (in the press).

ACKNOWLEDGEMENTS. We thank M. Sprinzl for discussion and communication of results before publication. G. Sawers for suggestions and reading the manuscript, and J. Auer for help with the sequence comparisons. This work was supported by a BMFT grant from the Munich Gene Centre.

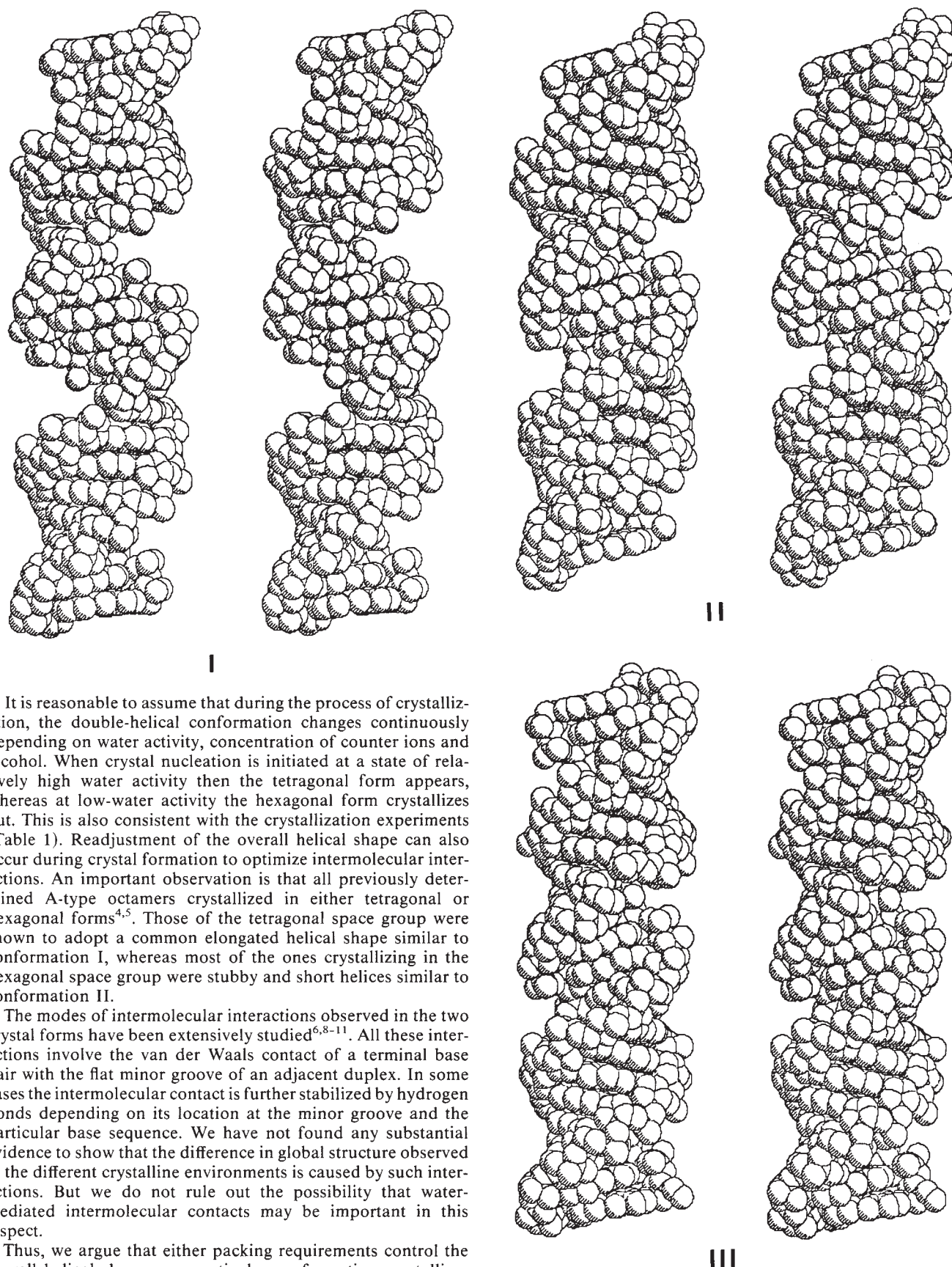


FIG. 1 Stereoscopic space-filling drawings of continuous helices based on conformations I, II and III. Each helix (24 base pairs long) was generated by rotating and translating the corresponding double-helical octamer with respect to the overall helix axis, using the average values of helix twist and rise per base pair of Table 2.

It is reasonable to assume that during the process of crystallization, the double-helical conformation changes continuously depending on water activity, concentration of counter ions and alcohol. When crystal nucleation is initiated at a state of relatively high water activity then the tetragonal form appears, whereas at low-water activity the hexagonal form crystallizes out. This is also consistent with the crystallization experiments (Table 1). Readjustment of the overall helical shape can also occur during crystal formation to optimize intermolecular interactions. An important observation is that all previously determined A-type octamers crystallized in either tetragonal or hexagonal forms^{4,5}. Those of the tetragonal space group were shown to adopt a common elongated helical shape similar to conformation I, whereas most of the ones crystallizing in the hexagonal space group were stubby and short helices similar to conformation II.

The modes of intermolecular interactions observed in the two crystal forms have been extensively studied^{6,8-11}. All these interactions involve the van der Waals contact of a terminal base pair with the flat minor groove of an adjacent duplex. In some cases the intermolecular contact is further stabilized by hydrogen bonds depending on its location at the minor groove and the particular base sequence. We have not found any substantial evidence to show that the difference in global structure observed in the different crystalline environments is caused by such interactions. But we do not rule out the possibility that water-mediated intermolecular contacts may be important in this respect.

Thus, we argue that either packing requirements control the overall helical shape or a particular conformation crystallizes in the appropriate crystal form. In any event, the important finding is that the observed structure and crystal packing are mutually dependent.

What is the role of the base sequence in determining the conformation of DNA in the crystalline state? Previous studies

have demonstrated a correlation between the local structure and the base sequence. The effect of the base sequence in A- and B-DNA is shown primarily by variations in base-stacking geometry and, to a much lesser extent, by changes in the sugar-phosphate backbone^{4,5}. Variations in the arrangement of the base pairs can be estimated from the relative rotations and translations of adjacent base pairs. The following three of these parameters are largely variable in the crystal structures: (1) the local helix twist, which is the relative rotation of two successive base pairs with respect to the helix axis; (2) the roll angle, a measure of the opening of the base pairs towards the helical grooves; and (3) the slide, which is the relative displacement of the base pairs along their long axes. The values of these parameters for each of the three structures are given in Table 2. Significant changes are observed from one crystal to the other.

The central pyrimidine-purine site of the sequence used in the study described here has been studied in several other A-type octamers: C-C-C-C-G-G-G-G (ref. 9), G-G-C-C-G-G-C-C (ref. 10, 11), and G-C-C-C-G-G-G-C (ref. 12). In all these structures as well as the tetragonal form of G-G-G-C-G-C-C-C (conformation I), the C-G step adopts a unique stacking geometry with a small twist (average 24°) and large negative slide (average -2 Å). The roll angles of the C-G doublets display a wide range of values (0-12°). The sugar-phosphate backbone of each of the C-G sites adopts an extended structure in contrast to the usual folded conformation observed in all other sites.

Only one of the three common properties is retained in the C-G step of conformation II, namely the helix twist (24°). The amount of slide is smaller (-1.3 Å). The values of the two parameters in conformation III (29° and -1.0 Å) differ much more from those of conformation I. The central backbones of

conformations II and III have the usual folded structure. The stacking geometries of the C-G doublets in conformations I, II and III are shown in Fig. 2. The change in sliding motion is reflected by the amount of overlapping between the central guanine bases. The overlapping diminishes gradually going from conformation I by way of II to III. The decrease in interstrand stacking between the guanines is associated with a simultaneous increase in intrastrand stacking between the guanine and cytosine bases.

The various parameters of several other helical steps also differ significantly between the different crystals. We believe that such local variations are not induced directly by crystal-packing forces for the following reasons. As mentioned before, the intermolecular interactions in all crystal forms are relatively soft and non-specific. Moreover, the 2-fold symmetry is largely retained in the duplexes that do not incorporate crystallographic symmetry axes^{1,4}. Even in cases where intermolecular interactions deform asymmetrically the double helix as for the B-dodecamer series², the local structure of equivalent base pairs is similar.

We suggest that changes in local structure of the same molecule in different crystalline environments are induced by hydration effects. This is supported by a common water structure observed in all C-G sites displaying the specific geometry described above (Fig. 3), in which a string of water molecules runs across the minor groove and interacts directly with the two guanines and two 3'-phosphate groups. This water structure is further stabilized by four water branches which are hydrogen-bonded to the sugar-phosphate backbone of two nearby symmetry-related duplexes. Any sliding of the base pairs in a direction that reduces the overlap of the guanines would disrupt this

TABLE 1 X-ray data of d(G-G-G-C-G-C-C-C)

	I	II	III
Temperature	293 K	293 K	100 K
Space group	$P4_32_12$	$P6_1$	$P6_1$
<i>a</i> (Å)	43.28	46.96	43.62
<i>b</i> (Å)	43.28	46.96	43.62
<i>c</i> (Å)	24.66	44.31	41.49
<i>V</i> (Å ³)	46,192	84,623	68,367
No. of oligomers	8	12	12
Resolution (Å)	1.7	2.9	1.9
No. of reflections at 2σ level	2,077	1,008	2,827
<i>R</i> -factor (%)	16.2	12.0	16.0
No. of ordered waters per duplex*	88	45	96

The deoxyoligonucleotide was prepared by the solid-phase procedure using fast phosphoramidite chemistry¹⁵, and purified by ion-exchange and HPLC. Crystallization was done at 19 °C using solutions containing 2 mM DNA, 5 mM MgCl₂, 0.15 mM spermine-HCl, 5% 2-methyl-2,4-pentanediol (MPD) and 50 mM sodium cacodylate (pH 7.0), equilibrated against reservoirs of concentrated MPD solutions (30 or 50%). Tetragonal crystals appeared predominantly at 30% MPD whereas hexagonal crystals appeared at 50% MPD when the drops almost dried out. The hexagonal crystals could be preserved only by adding buffered solutions containing 75% MPD. Diffraction data were measured at both RT and LT on a Rigaku AFC5 diffractometer equipped with a rotating anode source. For the RT experiment the crystal was sealed in a glass capillary containing a droplet of mother liquor. For the LT experiment a naked crystal was mounted directly and shock-cooled to 100 K (ref. 16). The structures were solved by the multi-dimensional search method ULTIMA¹⁷ and refined using the restrained least-squares program by Hendrickson and Konnert¹⁸. Solvent peaks located from electron density maps and satisfying hydrogen-bonding criterions were included in the refinement as oxygen atoms.

* Although the total number of ordered waters included in the refinement of conformation I was smaller than that of conformation III, the number of waters located at the first and second hydration shells of conformation I was 20 more than that for conformation III.

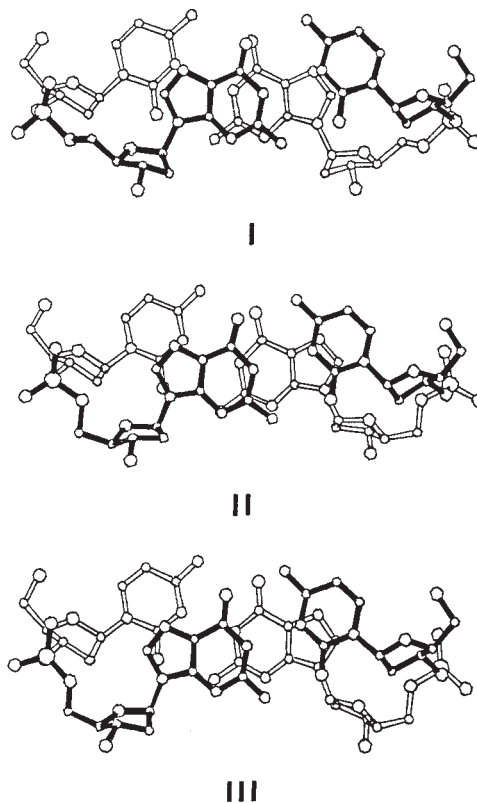


FIG. 2 Stacking geometry of C-G doublets in conformations I, II and III. The view is perpendicular to the upper base pair. It shows the decrease in slide from conformation I by way of II to III, and the increase in helix twist from conformations I and II to III.

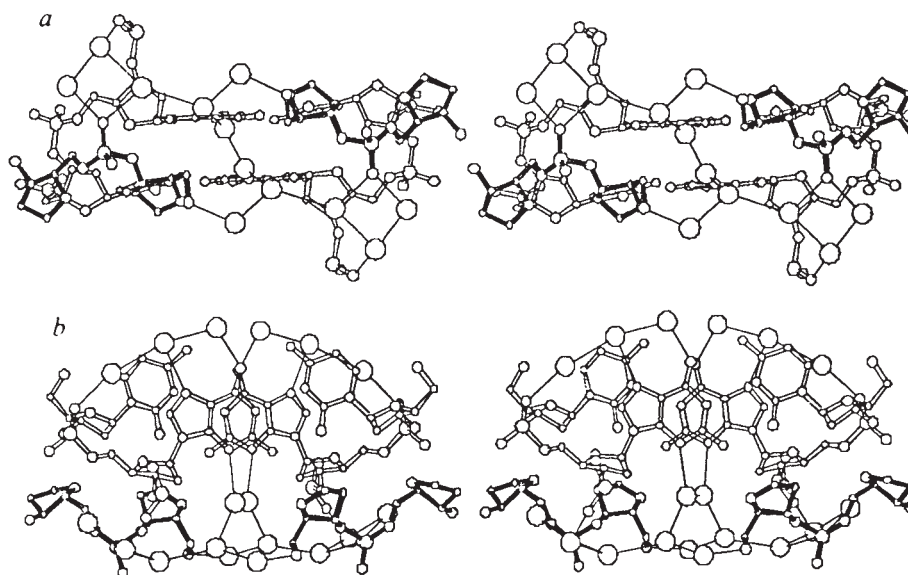


FIG. 3 Stereoscopic drawings of hydration of the C-G doublet in conformation I. *a*, The view is down the minor groove showing the interactions of water molecules with N3 atoms of guanines and the adjacent 3'-phosphate oxygens. The water structure is stabilized by hydrogen bonds to the sugar-

phosphate backbone of G1 and G2 from neighbouring duplexes (dark bonds). *b*, The view is perpendicular to the upper base pair, showing the hydration at both grooves. At the major groove a string of waters is hydrogen bonded to O6, N7 and 5'-phosphate oxygen atoms of each of the guanine residues.

hydration. At the major groove, strings of water molecules interact with the guanine bases and the 5'-phosphate oxygens of the locally extended backbone. Hence, the unique local structure observed in all the tetragonal C-G sites is stabilized by this water structure. An equivalent type of hydration cannot be accommodated in the hexagonal crystals because part of the central region of the minor groove is blocked by a neighbouring

molecule, thus demonstrating the indirect influence of crystal packing on local DNA structure.

The results presented here have implications on the modes of DNA recognition by proteins^{13,14}. The specificity of the recognition of a sequence of bases and hence the stability of the DNA-protein interface are dependent on the correct alignment of the base pairs along the helix. We show that a specific alignment of two successive base pairs is driven by a local hydration. It is therefore reasonable to anticipate that even larger changes in the orientation and position of the DNA base pairs can be accomplished by direct or water-mediated interactions with the protein. Also, specific hydration of the base pairs at one groove may govern the degree of helical twist, roll and slide needed to accommodate protein-DNA contacts at the other groove.

From the results of this study we conclude that the DNA double helix can adopt a range of conformations in the crystal-line state depending on hydration, molecular packing and temperature. Such studies demonstrate the capability of DNA to undergo important conformational changes as a result of modest variations in its environment. □

Note added in proof. Two conformations in the crystal were recently reported for another A-DNA octamer²¹.

TABLE 2 Helix parameters

Structure	Global parameters								
	I			II			III		
Helix twist (°)	32 (4)			31 (5)			32 (2)		
Rise per base-pair (Å)	3.3 (0.1)			2.9 (0.3)			2.9 (0.3)		
Base-pair inclination (°)	7 (1)			13 (2)			14 (1)		
Propeller twist (°)	-9 (3)			-6 (6)			-9 (5)		
Base-pair displacement dx (Å)	-3.8 (0.2)			-4.9 (0.3)			-3.5 (0.2)		
Helix diameter (Å)	19.5 (0.9)			19.7 (0.8)			18.1 (0.4)		
Major-groove width (Å)	10.9 (0.7)			8.0 (0.4)			5.6 (0.5)		
Minor-groove width (Å)	9.4 (0.3)			9.7 (0.6)			10.3 (0.5)		
Helical step	Local parameters								
	Twist (°)			Roll (°)			Slide (Å)		
	I	II	III	I	II	III	I	II	III
G-G	32.3	29.4	32.0	7.2	1.2	3.7	-1.4	-2.2	-1.6
G-G	31.2	28.9	36.0	6.4	13.2	5.5	-1.7	-2.2	-1.3
G-C	34.7	35.5	32.6	3.7	2.0	7.1	-1.3	-1.6	-0.7
C-G	24.0	23.7	29.0	3.5	8.9	10.0	-2.2	-1.3	-1.0
G-C	34.7	37.3	32.0	3.7	3.5	6.1	-1.3	-2.0	-0.8
C-C	31.2	29.9	33.4	6.4	10.3	7.5	-1.7	-2.4	-1.7
C-C	32.3	30.0	30.8	7.2	11.1	6.9	-1.4	-1.5	-1.5
Average	31.5	30.6	32.3	5.4	7.2	6.7	-1.6	-1.9	-1.2
s.d.	3.6	4.5	2.2	1.7	4.9	1.9	0.3	0.4	0.6

The numbering of residues from 5' to 3' is G1 to C8 in the first strand and G9 to C16 in the second. Hence, the first G-G step consists of base pairs G1-C16 and G2-C15, and so on. Helix diameter was calculated from the average distance of phosphorus atoms to the overall helix axis. The groove widths were estimated from averaging the shortest P-P distances across each groove of the continuous helix. Definitions of the other parameters are given in refs 2 and 19. The convention of signs used here is the one agreed on at an EMBO workshop on DNA curvature and bending²⁰. Hence, propeller twist, base-pair displacement and slide are negative unlike those of previous publications. The s.d. (given in parentheses) is the r.m.s. deviation from the mean. I, Tetragonal conformation; II, hexagonal RT conformation; III, hexagonal LT conformation.

Received 1 June; accepted 27 September 1989.

- Shakked, Z. & Kennard, O. in *Biological Macromolecules and Assemblies: Nucleic Acids and Interactive Proteins* Vol. 2 (eds Jurnak, F. A. & McPherson, A.) 1-36. (Wiley, New York, 1985).
- Dickerson, R. E., Kopka, K. L. & Pjura, P. in *Biological Macromolecules and Assemblies: Nucleic Acids and Interactive Proteins* Vol. 2 (eds Jurnak, F. A. & McPherson, A.) 38-126 (Wiley, New York, 1985).
- Wang, A. H.-J. & Rich, A. in *Biological Macromolecules and Assemblies: Nucleic Acids and Interactive Proteins* Vol. 2 (eds Jurnak, F. A. & McPherson, A.) 128-170 (Wiley, New York, 1985).
- Shakked, Z. & Rabinovich, D. *Progr. Biophys. molec. Biol.* **47**, 159-195 (1986).
- Haran, T. E. & Shakked, Z. *J. molec. Struct.* **179**, 367-391 (1988).
- Rabinovich, D., Haran, T. E., Eisenstein, M. & Shakked, Z. *J. molec. Biol.* **200**, 151-161 (1988).
- Eisenstein, M. *et al. Acta crystallogr.* **B44**, 625-628 (1988).
- Shakked, Z. *et al. Proc. R. Soc. B213*, 479-487 (1981).
- Haran, T. E., Shakked, Z., Wang, A., & Rich, A. *J. biomolec. Struct. Dynam.* **5**, 199-217 (1987).
- Wang, A. H.-J., Fujii, S., van Boom, J. H. & Rich, A. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3968-3972 (1982).
- Haran, T. E. thesis, Feinberg Grad. Sch. Weizmann Inst. Sci. (1986).
- Heinemann, U., Lauble, H., Frank, R. & Blocker, H. *Nucleic Acids Res.*
- von Hippel, P. H. & Berg, O. G. *Proc. natn. Acad. Sci. U.S.A.* **83**, 1608-1612 (1986).
- Otwinowski, Z. *et al. Nature* **335**, 321-329 (1988).
- Beaucage, S. L. & Caruthers, M. H. *Tetrahedron Lett.* **2**, 1859-1862 (1981).
- Hope, H. *Acta Crystallogr.* **B44**, 22-26 (1988).
- Rabinovich, D. & Shakked, Z. *Acta crystallogr.* **A40**, 195-200 (1984).

18. Hendrickson, W. A. & Konnert, J. H. in *Biomolecular Structure, Conformation, Function and Evolution* Vol. 1 (eds Srinivasan, R. Subramanian, E. & Yatindra, N.) 43-57 (Pergamon, Oxford, 1981).
 19. Fratini, A. V., Kopka, M. L., Drew, H. R. & Dickerson, R. E. *J. biol. Chem.* **257**, 14686-14707 (1982).
 20. Dickerson, R. E. *et al. EMBO J.* **8**, 1-4 (1989).
 21. Jain, S. & Sundaralingam, M. *J. J. biol. Chem.* **264**, 12780-12784 (1989).

ACKNOWLEDGEMENTS. This work was first presented at the EMBO workshop on DNA curvature and bending (Cambridge, September 1988). We thank the Minerva Foundation, Munich, Germany, and the US/Israel Binational Science Foundation for financial support. Z.S. holds the Helena Rubinstein Professorial chair in Structural Biology.

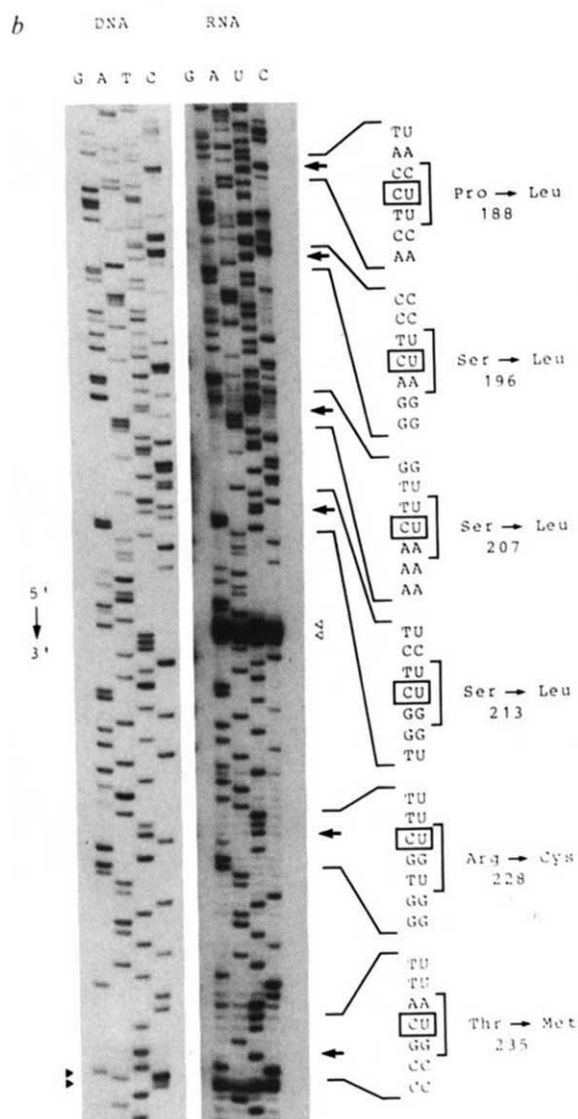
ERRATUM

RNA editing in plant mitochondria

Patrick S. Covello & Michael W. Gray

Nature **341**, 662-666 (1989)

FIGURE 2b in this letter was printed incorrectly, so that the gels were rotated through 180° and do not correspond to the labels. The correct version is printed below and will appear in reprints.



nature
is available in
microform.



University Microfilms

International reproduces this publication in microform: microfiche and 16mm or 35mm film. For information about this publication or any of the more than 13,000 titles we offer, complete and mail the coupon to: University Microfilms International, 300 N. Zeeb Road, Ann Arbor, MI 48106. Call us toll-free for an immediate response: 800-521-3044. Or call collect in Michigan, Alaska and Hawaii: 313-761-4700.

☐ Please send information about these titles:

Name _____

Company/Institution _____

Address _____

City _____

State _____ Zip _____

Phone () _____

University
Microfilms
International